

## THESIS / THÈSE

### MASTER EN SCIENCES INFORMATIQUES

#### Les interfaces multimodales dans les véhicules

Minne, Denis

*Award date:*  
2021

*Awarding institution:*  
Université de Namur

[Link to publication](#)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**UNIVERSITÉ  
DE NAMUR**

---

FACULTÉ  
D'INFORMATIQUE

## **Les interfaces multimodales dans les véhicules**

Denis Minne



# Table des matières

<b>1</b>	<b>Résumé</b>	<b>6</b>
<b>2</b>	<b>Remerciements</b>	<b>7</b>
<b>3</b>	<b>Introduction</b>	<b>8</b>
<b>4</b>	<b>Méthodologie de recherche</b>	<b>11</b>
4.1	Recherche . . . . .	11
4.2	Critères de sélection des documents . . . . .	11
<b>5</b>	<b>Design des interfaces</b>	<b>13</b>
5.1	Rôles et objectifs . . . . .	13
5.2	L'environnement dans le véhicule . . . . .	16
5.3	Les voitures autonomes . . . . .	17
5.4	Etude de ces technologies . . . . .	18
5.4.1	Evaluation objective . . . . .	19
5.4.2	Evaluation subjective . . . . .	19
5.4.3	Tests sur simulateur en laboratoire . . . . .	20
5.4.4	Sur route . . . . .	20
<b>6</b>	<b>Les interfaces d'entrée</b>	<b>22</b>
6.1	Types d'entrée . . . . .	22
6.2	Boutons . . . . .	23
6.2.1	Pression . . . . .	23
6.3	Touchscreen . . . . .	23
6.4	Gestuelle . . . . .	25
6.4.1	Technologies employées . . . . .	25
6.4.2	Techniques de reconnaissance . . . . .	26
6.4.3	Gestes possibles . . . . .	26
6.5	Reconnaissance vocale . . . . .	30
6.5.1	Techniques de reconnaissance vocale utilisées . . . . .	31
6.6	Regard . . . . .	31
6.7	Roulette . . . . .	32
6.8	Synthèse . . . . .	32



<b>7</b>	<b>Les interfaces de sortie</b>	<b>33</b>
7.1	L'écran - Touchscreen . . . . .	33
7.2	Système à vibration . . . . .	34
7.3	Utilisation de solénoïdes . . . . .	34
7.4	Système ultrasonique . . . . .	34
7.5	Système Sonore . . . . .	35
7.5.1	Spearcons . . . . .	36
7.5.2	Sonification . . . . .	36
7.5.3	Voix . . . . .	36
7.6	Afficheur frontal HUD (Head Up Display) . . . . .	38
7.6.1	Techniques . . . . .	38
7.6.2	Affichage d'informations sur le HUD . . . . .	39
7.7	Bandes de leds dans le champ visuel du conducteur . . . . .	40
7.8	Température . . . . .	41
7.9	Transmission d'odeurs . . . . .	41
7.10	Synthèse . . . . .	42
<b>8</b>	<b>Utilisation multimodale de ces interfaces</b>	<b>43</b>
8.1	Combinaison de modalités . . . . .	43
8.2	Interfaces multimodales en entrée . . . . .	45
8.3	Interfaces multimodales en sortie . . . . .	46
8.3.1	feedback . . . . .	47
<b>9</b>	<b>Problématisation</b>	<b>49</b>
<b>10</b>	<b>Etude d'un système de reconnaissance de gestuelle pour voi-</b>	<b>53</b>
	<b>ture</b>	
10.1	Actions proposées par le système de reconnaissance . . . . .	54
10.2	Gestuelle choisie . . . . .	55
10.3	Confirmation du geste . . . . .	58
<b>11</b>	<b>Implémentation</b>	<b>59</b>
11.1	Choix du matériel et des logiciels . . . . .	59
11.2	Interface de reconnaissance utilisée . . . . .	59
11.2.1	Performances du module Leap Motion . . . . .	61
11.3	Techniques utilisées pour la reconnaissance . . . . .	61
11.4	Différences entre main gauche et main droite . . . . .	64
11.5	Stockage des valeurs . . . . .	64
11.6	Implémentation du logiciel . . . . .	65
11.7	Feedback sonore au conducteur . . . . .	67
11.8	Evaluation et test du système . . . . .	68
11.9	Efficacité et précision du système. . . . .	70
11.9.1	Module "Leap Motion" . . . . .	70
11.10	Conséquences des problèmes de tolérances . . . . .	73

11.11Sécurité . . . . .	73
<b>12 Conclusion</b>	<b>74</b>
12.1 Etat de l'art . . . . .	74
12.2 Recherche . . . . .	75
<b>13 Recherches futures</b>	<b>77</b>
13.1 Interface . . . . .	77
13.2 Gestuelle . . . . .	77
<b>Glossaire</b>	<b>80</b>
<b>Bibliographie</b>	<b>82</b>

# Chapitre 1

## Résumé

Dans ce document, nous proposons de faire une étude des différentes interfaces de communication entre le conducteur et le véhicule. Il est possible qu'elles puissent également être utilisées par les passagers. Nous étudierons en particulier les interfaces à l'intérieur du véhicule permettant une communication volontaire, la façon dont elles ont été conçues et l'objectif recherché lors de leurs conception. Un élément central de cette conception pour beaucoup d'interfaces est la sécurité. Cet élément sera également très important dans ce document.

Dans un deuxième temps, nous proposons de faire l'étude d'une amélioration possible pour un type d'interface en particulier : les interfaces permettant la gestuelle en l'air pour commander le véhicule. Nous proposons d'étudier la possibilité de commander différents éléments du véhicule avec les deux mains. Est-ce possible ? Quelle amélioration cela apporte-t-il ? Quel est l'impact sur la sécurité ?

Mots clefs : voiture, gestuelle, interface multimodale

## Chapitre 2

# Remerciements

Nous tenons à remercier ici :

- Le Professeur Bruno Dumas, Professeur à l'Université de Namur, qui fût mon promoteur pour ce travail, et m'a fourni une aide indispensable.
- Le Vice-Doyen de la faculté d'informatique, Monsieur Laurent Schumacher pour m'avoir écouté et conseillé dans mon choix de sujet pour la réalisation de ce mémoire
- Le Professeur Pierre-Yves Schobbens, Professeur à l'Université de Namur, qui fût mon professeur pour l'initiation à la démarche scientifique.

D'un point de vue plus général, nous adressons nos plus vifs remerciements à l'ensemble des professeurs et assistants de l'université de Namur pour leur efficacité, la qualité exceptionnelle des cours qui m'ont été dispensés, leur compétence et leur accueil. Leurs encouragements nous ont donné la possibilité de nous relancer dans une phase d'amélioration et de remise à niveau de nos compétences.

## Chapitre 3

# Introduction

Depuis de nombreuses années, les constructeurs automobiles ajoutent à leurs véhicules bon nombre d'options et interfaces pour faciliter la vie du conducteur : air conditionné, GPS intégré, musique, radio, téléphone mains libres, accès à internet, aides à la conduite et bien d'autres choses encore... "Mainstream car manufacturers now offer up to 700 infotainment and environmental controls for the driver and passengers to manipulate" [47]

Toutes ces nouvelles options nécessitent d'ajouter des interfaces de communication dans le véhicule permettant au conducteur d'obtenir ce qu'il veut. Mais commander ces systèmes, bien que simples en apparence, pose un énorme problème : distraire le conducteur de sa tâche principale qu'est la conduite de son véhicule. De plus, celui-ci ne peut pas forcément s'arrêter pour pouvoir effectuer une action complexe ou demander de l'aide d'un passager. Il se retrouve donc à faire ses manipulations sans prendre conscience que son véhicule lancé à une vitesses de 120km/h parcourt 33m par seconde alors qu'il peut rester plusieurs secondes sans regarder la route.

On en arrive à un constat effrayant aujourd'hui : bon nombre d'accidents sont provoqués par l'inattention du conducteur qui ne remarque pas le danger alors qu'il manipule un élément tel que l'interface d'infotainment sur son tableau de bord. Celui-ci est distrait de sa première tâche, la conduite, au profit de sa tâche secondaire avec les conséquences dramatiques qu'on peut imaginer. "engaging in secondary tasks can reduce attention on driving and thus may affect road safety" [18] "Distractions from interactions with the radio or cell phones can lead to degraded driving performance across different age groups or road complexities and are estimated to cause 25% of vehicle crashes in the US" [18]

"At the dawn of the age of autonomous vehicles, alarming reports are showing that every year approximately 26,000 people die, and 1.4 million people

are injured in traffic accidents all around Europe. The great majority of these road accidents (about 90%) involve human error. Specifically, distracted driving in the USA accounts for approximately 25% of all motor vehicle crash fatalities, thus contributing to 3,477 traffic deaths in 2015 alone”[55]

Pour des réactions plus proactives du véhicule, les constructeurs se sont lancés également dans la construction de détecteurs permettant soit d’avertir le conducteur d’un accident potentiel ou d’une erreur qu’il aurait commise, soit intervenir sur la conduite du véhicule pour éviter l’accident[37]. Même si de telles technologies ont pour but d’améliorer la conduite du véhicule et donc la sécurité de celui-ci, elles posent le problème de la façon d’avertir le conducteur sans gêner sa conduite. Si un tel système attire l’attention sur un véhicule qui semble sur trajectoire mais en réalité garé sur le bas-côté, cela peut être contre-productif si, ce faisant, il détourne l’attention du conducteur qui ne remarque pas un enfant sur le point de traverser un peu plus loin.

Les constructeurs automobiles et les chercheurs rivalisent d’ingéniosité pour rendre les véhicules plus sûrs. Mais il est clair que l’élément principal de la sécurité des véhicules qu’ils construisent doit rester le conducteur, du moins tant que ces véhicules ne seront pas devenus des véhicules complètement autonomes. A partir du moment où le conducteur a un rôle à jouer dans la conduite, il est vital de s’assurer qu’il ne soit pas distrait sans raison.

Aujourd’hui, bien conscients de ce problème, beaucoup de constructeurs se sont lancés dans des recherches pour proposer de nouvelles interfaces de communication homme-machine permettant de diminuer le risque d’accident. Ces recherches ont déjà mené à bon nombre d’innovations ayant pour but de permettre au conducteur d’effectuer des tâches secondaires comme la manipulation de l’infotainment en limitant au maximum l’impact sur sa conduite. Cet impact peut être de deux types :

1. mental : si l’interface demande au conducteur trop de charge cognitive, il ne peut plus se concentrer sur sa conduite
2. sensoriel : un des exemples les plus grands est le fait que le conducteur doive regarder l’interface et ne peut donc plus regarder la route.

”Estimations show that it takes up to 1.1sec for a driver to glance from the road to the cluster and back on the road, and 0.6sec to glance from the road to the virtual image and back to the road.”[15]

En première partie, ce document passe en revue les solutions proposées par les chercheurs ou les constructeurs automobiles, celles mises au point et déjà produites en masse dans certains véhicules actuels ou celles ayant

simplement fait l'objet d'une étude. La partie principale de l'état de l'art présentera donc les différentes techniques : d'abord les techniques de communication d'informations par la voiture au conducteur, ensuite les techniques permettant au conducteur de commander le véhicule. Il est donc proposé dans ce document de procéder d'abord à une revue de ce qui existe déjà ou de ce qui a été proposé et étudié en ce qui concerne les interfaces dans les véhicules. On proposera d'abord une description de l'environnement et des technologies employées. On y expliquera les technologies employées et comment s'opèrent les choix et les évaluations de ces technologies. Ensuite viendra une revue des technologies existantes, permettant au conducteur de communiquer une information au véhicule, puis les technologies permettant de communiquer des informations au conducteur.

Sur base de cet état des lieux, nous proposons le développement d'une interface de communication homme-machine qui sera présentée plus loin dans ce document. Cette interface devra permettre de commander certains équipements du véhicule tout en maximisant la sécurité.

La logique de sécurité sera choisie comme élément central car la grande majorité des études de ces interfaces cherchent à connaître leur impact sur la conduite d'un véhicule.

## Chapitre 4

# Méthodologie de recherche

### 4.1 Recherche

Les sources principales de mes recherches sont :

- ACM Digital Library
- IEEE Xplore Digital Library

La recherche doit contenir "car" ou "cars" dans le titre. De plus les termes "multimodal interface", "speech", "gesture" doivent être proposés n'importe où dans le document. La recherche se limite à janvier 2016 pour les documents les plus anciens. Une première recherche sur DL.ACM.ORG a donné 131 documents.

Une autre recherche a ensuite été lancée sur [www.ieee.org](http://www.ieee.org). Elle a abouti à 42 documents.

Ensuite, un tri des documents sur base du résumé proposé a permis d'en écarter un grand nombre. D'autres ont fait l'objet d'une étude plus poussée car le résumé semblait prometteur mais au final, ils se sont révélés trop éloignés ou n'apportaient rien d'intéressant. Enfin, au cours des recherches, le cadre de ce mémoire a été revu à plusieurs reprises. Certains documents qui semblaient intéressants ont été rejetés, car les nouvelles limitations les rendaient hors sujet.

Au final, il reste un total 54 documents exploitables pour les deux recherches.

### 4.2 Critères de sélection des documents

Un nombre assez important de documents a été rejeté. Le premier tri avait pour but d'éliminer les documents qui ne parlaient pas de véhicules ou d'interfaces homme-machine. Par exemple, un document nommé 'CARS' était l'acronyme d'un sujet qui n'avait rien à voir avec les voitures. D'autres documents proposaient des études dans les véhicules, mais ces études se focalisaient sur d'autres aspects de ceux-ci. Dans quelques cas, ils parlaient



des interfaces mais ne proposaient pas de contenu qui pouvait donner des informations intéressantes pour notre sujet.

Ensuite, un grand nombre de documents avaient pour but d'étudier des capteurs dans le véhicule (détecteurs d'attention, de stress, etc...). Ces documents ont également été rejetés. Ce travail se limite aux technologies visant une communication volontaire entre le véhicule, le conducteur et/ou ses passager(s). Les systèmes de perception de stress, d'attention, de sommeil, ou tout autre capteur ne sont pas pris en compte et sont systématiquement rejetés.

Il a fallu revoir le cadre des recherches vu la grande quantité de documents. La question étant "quoi rejeter" ? Le choix suivant s'est imposé : seule la communication du conducteur et/ou des passagers du véhicules est prise en compte. Tous les documents proposant des interfaces extérieures pour les passants, les interactions extérieures avec le véhicule ne sont pas pris en compte. Notons que certains documents proposent une communication pouvant se faire à l'intérieur avec le conducteur et à l'extérieur du véhicule avec l'entourage. Dans ce cas, le document est pris en compte pour ce qui est de la communication avec le conducteur à l'intérieur du véhicule. Le reste est rejeté.

La prise en compte des passagers à l'intérieur du véhicule a également été soumise à réflexion mais il a été décidé de les inclure pour peu qu'ils se trouvent réellement à l'intérieur du véhicule. Enfin, un dernier problème était celui des véhicules autonomes. Initialement, il a été prévu de rejeter les documents qui en parlaient, mais après étude des documents, la question s'est révélée plus complexe que prévu. Il apparaît que les véhicules autonomes changent fortement la logique de design des interfaces de communication homme-machine. Cet élément ne peut donc être ignoré dans notre recherche, bien qu'il ne constitue pas un élément central.

## Chapitre 5

# Design des interfaces

### 5.1 Rôles et objectifs

Tout d'abord, il faut définir ce que sont les interfaces multimodales dans les véhicules. Les interfaces multimodales sont des interfaces qu'on va utiliser en les regroupant avec d'autres interfaces pour permettre une communication complexe. La communication entre le conducteur et le véhicule passe par plusieurs phases (Fig. 5.1) :

1. Décision : le conducteur décide d'effectuer une action.
2. Action : le conducteur transmet sa demande au véhicule via un moyen de communication à sa disposition.
3. Perception : le véhicule reçoit via une ou plusieurs interfaces de communication.
4. Interprétation : l'ordinateur interprète les informations reçues des différentes interfaces, en déduit la demande et effectue la ou les actions demandées.

La procédure inverse se déroule lorsque le véhicule veut transmettre une information au conducteur. Le véhicule veut transmettre une information et va activer une ou plusieurs interfaces de communication. Le conducteur reçoit l'information via l'un de ses sens (ouïe, toucher, vue, ...). Il va enfin traiter les informations reçues et en déduire l'information transmise.

Mais dans un véhicule, le conducteur doit réserver son attention à la conduite qui est son activité principale. Cette activité peut déjà être considérée comme une activité multimodale dans le sens où elle accapare à elle seule presque tous les sens du conducteur. Même si certains signaux transmis du véhicule ou transmis au véhicule ont pour but une aide à la conduite, il faut considérer que la plupart de ces signaux sont exploités pour une activité secondaire. Il faut donc penser la communication avec le conducteur en prenant cet élément comme base de travail.

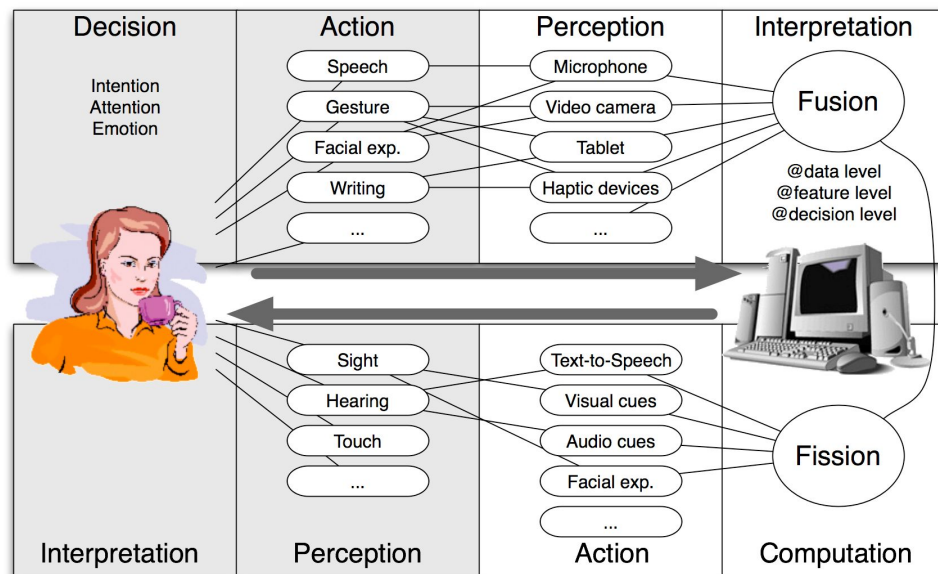


FIGURE 5.1 – source : Multimodal Interfaces : A Survey of Principles, Models and Frameworks - Bruno Dumas, Denis Lalanne, Sharon Oviatt

Beaucoup d'accidents aujourd'hui ont pour cause la distraction du conducteur[18]. Bien conscients du problème, les chercheurs et constructeurs cherchent à développer de nouvelles interfaces en prenant en compte ce fait. Les recherches visant à développer ces interfaces ont pour but premier la sécurité. L'idée pour cela est d'évaluer l'interface en tenant compte du fait que le conducteur puisse l'utiliser en gardant autant que possible les yeux sur la route, les mains sur le volant et qu'il puisse rester concentré sur sa conduite. En résumé, il est important que la tâche première qu'est la conduite n'interfère pas sur sa seconde tâche. Les études d'interfaces ont donc pour but de comparer l'avantage ou la perte d'attention sur la conduite qu'engendre cette nouvelle interface par rapport à une ou plusieurs autres. Si une interface se montre trop gourmande sur les performances de conduite, elle perd son intérêt. L'idée est donc de permettre au conducteur de se focaliser sur sa tâche première et optimiser la seconde pour diminuer l'attention nécessaire et la distraction provoquée avec entre autres l'approche "les yeux sur la route" [2]

Les interfaces dans les véhicules aujourd'hui peuvent avoir plusieurs rôles :

1. gestion de la conduite : mise en marche des essuie-glace avant et arrière et modification de leur vitesse, clignotants, allumage et réglage des phares, ...

2. information sur la conduite : données de vitesse, avertissement de risque d'accident, distances de sécurité, angle mort, franchissement de ligne blanche, ...
3. infotainment : radio, musique, GPS, téléphonie, ...
4. environnement du véhicule : ouverture ou fermeture des vitres, chauffage, air-conditionné, réglage des sièges, ...
5. le monitoring du véhicule.

Le choix et le positionnement de ces interfaces doit être étudié. Le conducteur doit pouvoir accéder au moment opportun à toutes ces interfaces sans que celles-ci ne puissent constituer une gêne : "The layout of cars is still constrained by traditional norms like the ISO 3958 :1996 which specifies the area a driver needs to reach with the hands to operate the vehicle "[54]

De plus le choix de l'interface de communication transmission et réception doit être choisie au mieux en fonction du message que l'on souhaite faire passer. Transmettre un avertissement de risque d'accident va plutôt se faire avec un signal sonore approprié pour indiquer l'urgence alors qu'une simple information de navigation exploitera un signal moins agressif. Pour ce point également, le type ou la complexité de l'action à effectuer a son mot à dire : demander en vocal "fermer la fenêtre un peu plus" n'est pas vraiment pratique pour avoir de la précision.[27].

Le document [9] pose la question "que souhaitez-vous faire dans votre véhicule ? " Les réponses montrent qu'en plus de la gestion du véhicule, les gens espèrent pouvoir y faire énormément d'activités mais certaines peuvent diminuer l'efficacité de l'activité principale qu'est la conduite. Il a été démontré que si l'interface utilisée est bien choisie, il est possible d'empêcher les performances de la conduite de baisser de façon critique [18].

Le problème législatif est aussi très important pour les constructeurs. Si certains pays ont une législation assez permissive en matière d'ajout d'interfaces dans le véhicule, d'autres exigent des validations et ont mis en place des règles très strictes. Produire un véhicule dont une interface n'est pas légale dans un certain nombre de pays oblige le constructeur à proposer, pour ces pays, une autre interface qui pourrait être moins efficace.

Une autre application de ces interfaces est le jeu vidéo. Il est clair qu'utiliser des jeux vidéos dans les véhicules peu sembler fortement accidentogène, mais si un ou plusieurs passagers du véhicule sont des enfants, ils peuvent accaparer l'attention du conducteur sans avoir conscience du danger que cela constitue. Contrairement à des passagers adultes et attentifs, ils ne sont pas forcément capables de réaliser que la situation du véhicule sur la route exige toute l'attention du conducteur. Ce risque est 12x plus important qu'en

utilisant un téléphone [25]. Mais il faut éviter que ces jeux ne puissent distraire le conducteur, ou gênent sa vue. L'article [20] propose l'utilisation des fenêtres interactives pour distraire les passagers pendant le voyage en exploitant l'environnement du véhicule, ou même via des interactions avec d'autres véhicules[14]. Mais on peut se demander si un tel système ne risque pas de diminuer la visibilité de l'angle mort pour le conducteur lors d'un dépassement par exemple.

Un problème de ces interfaces est qu'elles peuvent nécessiter des configurations qui seront propres au conducteur, tout comme le sont les sièges du véhicule. Une solution serait justement de reconnaître le conducteur pour lui remettre sa configuration en place dès qu'il se place au volant. Cela permettrait d'éviter une reconfiguration fastidieuse des interfaces et accessoires du véhicule. Le document [49] propose pour cela l'utilisation de la voix, mais nous le verrons un peu plus loin, cette option n'est pas forcément judicieuse car l'environnement direct du véhicule est peut-être trop bruyant pour le permettre. Dès lors, proposer une autre interface permettrait d'y remédier, même si elle est plus complexe à utiliser.

## 5.2 L'environnement dans le véhicule

L'utilisation de telles interfaces semble au premier abord une bonne idée, mais il ne faut pas oublier qu'une voiture doit évoluer dans son environnement. Si les constructeurs ont fait de gros efforts pour améliorer l'habitacle du véhicule où de telles interfaces sont utilisées, cet environnement reste très difficile à gérer. Nous le verrons plus tard mais peu d'interfaces sont épargnées par ces problèmes. Certaines technologies ou techniques utilisées dans d'autres secteurs sont dans ce cas inapplicables : identifier une personne par sa démarche par exemple [34]

Pour les interfaces sonores, les bruits extérieurs au véhicule, le bruit constant du moteur, la musique, le bruit provoqué par les fenêtres ouvertes ou les bruits de roulement peuvent rendre un signal sonore ou une interface inaudible pour le conducteur. Pour limiter ce problème, on peut proposer d'utiliser l'effet Lombart pour rendre le son plus audible sans pour autant augmenter le volume[26]. Enfin, émettre des sons ou voix peut se révéler gênant pour le conducteur dès lors qu'une conversation est en cours avec un passager ou via son GSM.

Pour les interfaces visuelles, ce n'est guère mieux : l'utilisation en plein soleil, en pleine nuit ou en présence de certains types d'éclairages extérieurs gênants, on peut aisément se retrouver dans une situation de pollution lumineuse importante. Il est possible que les signaux de l'interface soient mal

perçus voire pas du tout.

Pour ce qui est des interfaces tactiles, le mauvais état de la route ou les dispositifs ralentisseurs peuvent avoir un impact négatif sur l'utilisation de l'interface[54]. La partie de gestuelle en l'air est très impactée par ce genre de problème au point d'empêcher leur utilisation, voire de provoquer des lectures incorrectes. "Road perturbations will continue to exert forces that can render in-vehicle touch interaction challenging." [22] "We show that road bumps have a significant effect on touch input and can decrease accuracy by 19%." [22] Ce document propose une technique pour diminuer ces erreurs de 32% avec du deep learning.

Si une interface se montre efficace dans un milieu classique, il ne faut pas oublier qu'une interface infrarouge par exemple indique une température de fonctionnement du style -20 à 40 °C. Dans certains pays, le véhicule peut se retrouver facilement dans des conditions hors de cette zone d'utilisation. Un véhicule en plein soleil peut facilement voir grimper sa température et ce type de technologies est alors rendu inopérant.

Enfin, si une interface peut se montrer problématique voire inefficace dans son environnement, son impact sur la sécurité peut être moindre voire contre-productif. Cependant, l'utilisation combinée de plusieurs technologies comme une caméra classique avec un système infrarouge ou tout autre système peut permettre de diminuer fortement ces problèmes[47].

## 5.3 Les voitures autonomes

Il existe plusieurs types de véhicules autonomes. Toutes les recherches ne sont pas d'accord sur les niveaux de conduite automatique. Le document [9] propose 5 niveaux de conduite autonome :

- 0 pas d'automatisme
- 1 fonctions spécifiques
- 2 fonctions combinées
- 3 conduite autonome limitée
- 4 conduite autonome totale

L'article [33] propose l'utilisation standard décrite dans le SAE J3016 qui se limite à 4 niveaux seulement. Cependant, l'idée est très proche : proposer différents niveaux d'autonomie, partant de rien, aucune autonomie et aboutissant à une autonomie complète sans conducteur.

Dans ce document, on s'intéresse presque exclusivement aux véhicules ne

disposant pas d'une technologie de conduite autonome suffisante pour permettre de se passer d'un conducteur. Dans ces conditions, les interfaces de communication homme-machine sont développées avec la sécurité comme élément central dans presque tous les cas et le fait que la conduite est l'activité principale du conducteur. Le conducteur doit donc autant que possible garder les yeux sur la route et les mains sur le volant pour rester concentré sur sa conduite. L'interface développée a donc pour but de limiter autant que possible les interactions et la distraction provoquées par son utilisation.

Dans le cas des véhicules proposant une technologie de conduite autonome limitée, un conducteur doit pouvoir reprendre le contrôle du véhicule à tout moment. Cela signifie que ce dernier doit proposer un environnement identique où presque à un véhicule sans gestion autonome. Dans le cas où le véhicule roulerait en configuration autonome, d'autres interfaces proposeraient une activité qui deviendrait l'activité principale à la place de la conduite : travail, relaxation, sport, ...[48] ou même l'utilisation de jeux entre les voitures[14]. Les interfaces existantes peuvent aussi se montrer beaucoup plus intrusives dans l'attention puisqu'il n'y a à ce moment plus de conducteur. On pourrait utiliser l'ensemble du pare-brise du véhicule comme un écran pour l'affichage d'applications [9]. Dans le cas spécifique de ces véhicules semi-autonomes, l'interface permettant la reprise de contrôle doit être pensée au mieux pour permettre une reprise de contrôle dans de bonnes conditions et empêcher une désactivation accidentelle : parole à répéter, décompte vocal ou réponse à une question posée[33].

Dans le cas des véhicules avec une technologie de conduite autonome totale, il n'y a plus besoin des interfaces conducteur. Lorsqu'on remarque que la grande quantité de documents scientifiques sur ces interfaces de communication homme-machine est très axée sur la problématique de la sécurité, on peut raisonnablement penser que ces véhicules vont fortement évoluer par rapport aux véhicules ne proposant au mieux qu'une technologie de conduite autonome limitée que nous connaissons aujourd'hui. Ces véhicules autonomes proposeraient donc des interfaces facilitant la communication des passagers [23] ou permettant au véhicule d'indiquer aux passagers quelles sont ses intentions pour rendre le voyage plus agréable [39].

## 5.4 Etude de ces technologies

Dans les document étudiés proposant de nouvelles interfaces ou une évaluation de différentes interfaces, pour procéder à une évaluation de leur efficacité, on procède en général à une comparaison avec une autre interface utilisée pour effectuer la même action. L'idée est de voir si cette nouvelle interface permet d'effectuer des actions avec plus de sécurité qu'avec une

interface déjà connue comme un touchscreen au centre du tableau de bord. Cette évaluation se fait en général de deux façons :

1. Une évaluation objective qui va évaluer le temps demandé ou l'impact sur les performances de conduite
2. Une évaluation subjective qui va évaluer par des questionnaires le ressenti des conducteurs en situation

[35, 55, 18, 45, 15, 33, 29, 28, 24, 21, 1, 44, 2, 3, 37, 40]

Ces deux tests peuvent donner des résultats contradictoires car une interface ou une utilisation de cette interface donne de meilleurs résultats en diminuant l'impact sur la qualité de conduite ou la vitesse d'action mais ne reçoit pas la préférence dans la comparaison objective dans les questionnaires[3, 29]. Cette évaluation peut se faire de deux façons différentes : en simulateur avec un matériel spécifique ou sur route en conditions réelles, plus rarement dans un environnement recréé. Un autre problème est que comparer une nouvelle technologie avec des technologies existantes déjà massivement utilisées peut provoquer des résultats discutables car les effets sur la conduite et le ressenti ne seront pas forcément les mêmes. Certaines études vont proposer au testeur de se familiariser avec l'interface, d'autres ne le font pas. On ne peut donc comparer les chiffres autrement que par les comparaisons faites directement dans les documents.

#### 5.4.1 Evaluation objective

Pour l'évaluation objective, on va mesurer :

1. l'impact sur la conduite le plus souvent via un simulateur en laboratoire : on mesure le temps de réaction, les fautes de conduite (déviation de trajectoire, excès de vitesse, ...), les fautes ayant pu engendrer un accident en condition réelle
2. l'efficacité de l'interface : le temps nécessaire pour évaluer une action et le taux d'erreur rencontré.

Ces deux éléments sont comparés avec d'autres techniques connues, par exemple : on compare la gestuelle en l'air avec l'utilisation d'un touchpad pour une même action. Les différences de résultats permettent d'évaluer si la nouvelle interface a montré de meilleurs résultats dans les conditions de tests.

#### 5.4.2 Evaluation subjective

L'évaluation subjective va permettre d'évaluer le ressenti. On utilise différents types de questionnaires pour évaluer des éléments tels que la charge mentale sur le conducteur ou ses préférences dans les différentes utilisations



qu'il a faites d'une interface. La plus fréquemment utilisée est la NASA-TLX [13, 3, 37, 44, 45, 1, 24, 2] qui permet d'évaluer :

1. la demande mentale : facile ou exigeante
2. la demande physique : quelle activité physique était nécessaire ?
3. la demande temporelle : avez-vous ressenti que cela demandait beaucoup de temps ? Avez-vous trouvé cela lent ou rapide ?
4. la performance : dans quelle mesure avez-vous réussi à accomplir la tâche ?
5. l'effort : à quel point avez-vous dû travailler mentalement et physiquement pour accomplir la tâche ?
6. la frustration : à quel point vous êtes vous senti frustré ou irrité pour réaliser cette tâche ?

D'autres tests pour évaluer l'interface sont proposés dans certains documents : SUS [40, 3], UEQ [36], WMT [24], WWL [24], CTAM [40], ATI [40], TA-EG [42], PXI [14], PANAVA-KS [2]

### 5.4.3 Tests sur simulateur en laboratoire

Dans la majorité des cas, les tests sont réalisés en laboratoire. On réalise le test avec un certain nombre de testeurs qui doivent effectuer des opérations sur un simulateur avec un volant et utiliser l'interface sur le simulateur. Il serait impossible de réaliser ces tests en milieu réel : modifier un véhicule peut être très dangereux et est interdit par la législation en vigueur dans de nombreux pays.

On pourrait effectuer les tests en milieu urbain simulé, mais il faut disposer de l'infrastructure adéquate. Dans ces conditions, au moins pour un premier essai, la solution est d'utiliser un simulateur et voir si, dans ce cas, l'interface se révèle déjà prometteuse. Mais pour prouver son efficacité, il faudrait passer par un test en conditions réelles dans un environnement beaucoup moins conciliant.

### 5.4.4 Sur route

Les tests sur route ne sont pas possibles pour toutes les interfaces. Seuls les constructeurs disposent des infrastructures adéquates pour procéder aux tests nécessaires en toute sécurité. Certaines interfaces peuvent être utilisées légalement dans les véhicules mais pas toutes. De plus les conditions de test diffèrent d'un pays à l'autre (état des routes, météo, ...).

On peut donc se demander si une interface ayant montré son efficacité sur une route précise dans une région montrera les mêmes résultats ailleurs à un autre moment.

Par exemple : [20] propose deux tests, un en ville et l'autre sur route rapide.

## Chapitre 6

# Les interfaces d'entrée

### 6.1 Types d'entrée

Aujourd'hui, dans les véhicules les techniques d'entrée ont pour but de permettre différentes actions : entrer une adresse de destination dans le GPS, rechercher une musique dans l'interface d'infotainment, régler un équipement, etc... Ces différentes actions doivent être accomplies assez rapidement et sans accaparer l'attention du conducteur mais elles peuvent nécessiter une logique d'entrée adaptée. Choisir une rue pour programmer le GPS serait fastidieux s'il fallait passer d'une rue à l'autre parmi toutes les rues du pays. "Many tasks such finding a song in a playlist or searching for a location in the GPS application may need significant scrolling actions and can take a considerable amount of time to complete. As a result, drivers are likely to be distracted from driving in a safe manner when interaction is required." [52].

C'est pour cela que l'interface d'infotainment pour un GPS par exemple propose un clavier d'entrée similaire au claviers d'ordinateurs, des listes de sélection rapide pour naviguer dans les rues ou dans les musiques disponibles et enfin des boutons pour confirmer ou effectuer d'autres actions. On en arrive donc à plusieurs techniques d'entrée :

1. boutons : bouton physique, gestuelle, bouton touchscreen, ...
2. scrolling : pression[29, 52], scrollbar sur touchscreen, roulette[29], gestuelle ...
3. texte : clavier[13], écriture[13], parole[13], ...

On peut donc en déduire qu'une interface d'entrée du véhicule peut être très efficace pour un type d'action mais ne le sera pas forcément pour d'autres types d'actions. Il faut donc évaluer l'utilité de l'interface au cas par cas. [13]

## 6.2 Boutons

Ce sont les plus anciens systèmes aujourd'hui. Bien que toujours utilisés pour l'activation des équipements de conduite, ils tendent à disparaître pour ce qui est de l'infotainment au profit des touchscreen. Dans les documents scientifiques, leur présence est surtout là pour permettre une comparaison avec de nouvelles interfaces.

Leur gros désavantage est que trop de boutons rend l'utilisation de l'infotainment trop complexe. Il faut donc se limiter à des boutons multifonctions sur le bord d'un écran LCD par exemple. Pourtant, cette solution montre encore aujourd'hui des avantages : le fait que les boutons ont une existence physique et permettent un retour tactile, un conducteur connaissant bien son véhicule peut aller chercher ce bouton des doigts sans avoir à quitter la route des yeux. Mais même s'il peut le faire, son utilisation en mode multifonctions l'oblige à regarder l'écran pour savoir quelle action il va déclencher.

### 6.2.1 Pression

Une évolution possible des boutons serait de leur permettre de ressentir la force émise lorsqu'on appuie dessus et de pouvoir exploiter cette information. Il a été démontré qu'on pouvait différencier six niveaux d'interaction sur une souris. [29]

En fonction de la force appliquée, le bouton peut par exemple servir à actionner l'un ou l'autre bouton "virtuel". On peut également s'en servir pour une sélection dans une liste [29].

Cependant, s'il a l'avantage de demander peu ou pas d'attention visuelle, il faut prévoir dans son développement et son utilisation le fait que l'environnement du véhicule en mouvement peut provoquer des activations intempestives si le système devait se montrer trop sensible. Le document [28] propose de ne pas reconnaître d'interaction en dessous de 3 newton.

Enfin, comparé à une roulette ou à une utilisation directe du touchscreen, ce système s'est montré le plus lent mais assez précis (94% de précision)[28].

## 6.3 Touchscreen

Les touchscreen sont très fréquemment employés dans les véhicules aujourd'hui pour l'utilisation de l'infotainment. Les recherches actuelles tendent cependant à ne plus utiliser cette technologie car on remarque qu'elle est l'une des plus demandeuses d'attention visuelle. Les touchscreens ont l'avantage de proposer une manipulation directe pour la communication. De plus,

ils permettent d'exécuter aisément les différents types d'action (bouton, sélection dans une liste et entrée de texte assez efficace). Son vrai problème reste le déficit d'attention qu'il provoque sur le conducteur, source d'accident.

Cependant, la façon de l'utiliser peut permettre de réduire ce défaut. On peut par exemple chercher à définir la taille de caractère la plus adaptée. [13] Une police de caractère plus petite augmente la quantité d'informations affichable à l'écran mais demande plus d'attention. Si cette police est plus grande, l'attention visuelle nécessaire est moindre, la sélection est plus facile et moins sujette à erreur.

Pour ce qui est de l'entrée de texte, on peut proposer deux façons possibles : l'utilisation d'un clavier, comme c'est très fréquent aujourd'hui ou une écriture à la main sur le touchscreen. [13] propose cette façon de faire. Mais cette technique a montré ses limites. Le comparatif proposé entre ces deux méthodes d'entrée a montré une certaine amélioration en ce qui concerne la perte d'attention de la route, mais est loin d'être suffisante pour convaincre. De plus les tests étant réalisés en laboratoire, rien ne prouve que les faibles résultats seront maintenus en conditions réelles.

L'utilisation classique des touchscreen (bouton, scrolling et clavier) est assez efficace mais ce système est assez sensible aux casse-vitesse, dos d'âne et autres ornières présentes sur les routes. Ceux-ci peuvent diminuer la précision de sélection de 19% et une augmentation du temps nécessaire à la sélection [22]. Cependant il est proposé d'utiliser du deep learning pour améliorer la précision.

Une technique intéressante pour les touchscreen peut être de pouvoir effectuer une sélection avec la main en l'air. [1] Le système permet de sélectionner l'élément voulu à l'écran en déplaçant la main mais sans devoir aller jusqu'à toucher l'écran. Un contrôleur de type Leap Motion permet de détecter la position et le geste de la main dans le véhicule et de sélectionner un point parmi 21 points de sélection possibles. Ici aussi, ce système a montré une forte diminution d'efficacité en cas d'utilisation sur une route dégradée.

Une autre technique possible pour améliorer son efficacité est de permettre de mesurer la force de pression lorsque le conducteur sélectionne un élément sur le touchpad ou via un capteur à proximité[10]. Cela permet une utilisation différente de l'interface qu'avec un système de boutons simples. On peut choisir deux types de pressions : 200 gr et 750 gr. Ces deux pressions permettent une perception et une réaction différentes par le véhicule.

## 6.4 Gestuelle

La gestuelle est une technologie prometteuse dans les véhicules. Elle permet au conducteur d'envoyer des commandes au véhicule sans avoir à quitter la route des yeux et de préserver en grande partie son attention. Son désavantage est la perte du feedback tactile. "A key advantage of mid-air HG interaction interfaces is their ability to reduce the mental and visual demand of the IVIS on the driver. Meanwhile, the primary disadvantage of pure HG interaction control is the loss of tactile feedback" [55]. En matière de gestuelle, on peut en distinguer deux types :

1. la gestuelle en l'air : cette gestuelle s'effectue en enlevant une main du volant et le geste doit être effectué dans une zone clairement définie dans l'habitacle du véhicule. Ce geste peut être effectué avec la main droite ou la main gauche en fonction de la position de la zone de détection. Cette gestuelle peut être une simple position de la main, un mouvement de celle-ci ou une combinaison des deux.
2. la micro-gestuelle : elle consiste à effectuer une action de la main mais en gardant celle-ci sur le volant autant que possible. Ici aussi, elle peut s'effectuer avec l'une ou l'autre main.

### 6.4.1 Technologies employées

L'utilisation de gestuelle dans l'environnement d'un véhicule est assez complexe. D'abord, et comme précisé précédemment, on ne peut positionner l'interface n'importe où. Son accès doit respecter la zone d'action du conducteur et sa position doit être aisément accessible. Le conducteur doit pouvoir y accéder sans avoir à effectuer de mouvement ni devoir la chercher du regard. De plus son positionnement doit respecter les contraintes de l'ISO3958 :1996 qui spécifie là où le conducteur peut mettre les mains[54]. Enfin, le système proposé doit être efficace, car là aussi, un système incapable de détecter et décoder correctement la demande serait contre-productif pour la sécurité. Même simple, un tel système nécessite un apprentissage des gestes proposés.[55]

L'habitacle d'un véhicule est fortement influencé par l'environnement et certaines interfaces se révèlent inadaptées aux variations de la luminosité ambiante[50]. C'est pour cela que différents systèmes sont proposés. Certains se montrent plus efficaces que d'autres. On peut utiliser une caméra infra-rouge[17, 51], une caméra time-of-flight[17, 53], un radar[17, 47], un dispositif "Kinect"[17] voire même un "gant" avec détection de mouvement via un accéléromètre. Dans le document [17], on propose d'utiliser simultanément la caméra infra-rouge, la caméra time-of-flight et le radar pour améliorer la perception. Cependant, si un système radar peut se montrer

plus efficace dans des conditions difficiles, il se montre souvent moins efficace pour la reconnaissance de certains types de gestes. Pour une utilisation de ce type d'interface, la gestuelle doit donc être pensée en conséquence.

### 6.4.2 Techniques de reconnaissance

Pour reconnaître les gestes proposés, on peut utiliser des techniques de réseaux neuronaux, soit pour reconnaître la position de la main, soit pour reconnaître le geste exécuté parmi une liste de gestes connus. La difficulté avec de telles techniques est qu'il faut pouvoir les entraîner suffisamment. Une reconnaissance en LSTM peut nécessiter 6000 échantillons[53]. A cela doit encore s'ajouter un certain nombre d'échantillons de tests.

Différentes techniques de réseaux neuronaux peuvent être utilisées comme le LSTM(Long short term memory)[53], le CNN (Convolutional Neural Network) [38] ou le DDNN (Distributed Deep Neural Network) [50].

### 6.4.3 Gestes possibles

"Recently, Angelini et al. presented a user-elicited taxonomy of gestures performed on the entire surface of a standard steering wheel. A larger variety of different gestures were proposed by the users and two different trends were evidenced : some people suggested gestures to be performed without taking the hands off the steering wheel, while others suggested gestures that required taking at least one hand off the steering wheel." [2] Afin de se faciliter la tâche, on donne généralement un nom aux différentes commandes possibles. De plus, dans les cas de gestes où la main reste correctement positionnée sur le volant, on parlera plutôt de micro-gestuelle. Les gestes proposés ici exigent d'enlever brièvement mais totalement les mains du volant, une à la fois et pendant un temps qu'on voudra le plus court possible.

Il existe une grande variété de gestes possibles. On peut les classer en deux catégories différentes : la position des doigts de la main d'une part et les mouvements de déplacement de la main d'autre part. Comme cela a été dit précédemment, les différentes technologies proposées ne permettent pas de reconnaître tous les gestes de façon identique : un système radar peut se montrer efficace pour reconnaître un déplacement de la main mais se montrer inefficace pour ce qui est de la position des doigts. La gestuelle choisie doit donc tenir compte de la technologie qui va la reconnaître. Ce n'est cependant pas la seule chose à prendre en compte. L'article [2] nous indique que le choix de la gestuelle a un impact sur l'intérêt perçu par les utilisateurs et la réponse émotionnelle. L'approche du choix de la gestuelle

doit être optimisée pour obtenir de bons résultats.

Plus la gestuelle est complexe, plus elle a de chances d'échouer. Il faut aller vers des gestes simples et courts[44]. Une gestuelle inadaptée va provoquer un grand nombre d'erreurs d'interaction des utilisateurs. Pour ce type de gestion un apprentissage peut être nécessaire. On conseillera donc, pour limiter la charge mentale de cette technique et simplifier son utilisation, de limiter au maximum les gestes proposés. Certains mouvements, tels qu'un geste circulaire, devront également être évités.[55]

Un autre élément important est le ressenti des utilisateurs vis-à-vis de ces gestes. D'abord certains gestes semblent plutôt présenter un ressenti positif et d'autres un ressenti négatif. Certains gestes seront donc plus adaptés pour un "Oui" et d'autre pour un "Non" [55]. Enfin un autre problème est qu'une utilisation excessive d'un tel système peut interférer avec la conduite et provoquer un effet de fatigue [55].

## Gestes

En ce qui concerne la position des doigts on peut par exemple proposer :

1. donner des valeurs de 1 à 5[38] et les contrôles Fist, flat, grab, pich, point (Fig. 6.1 & 6.2).
2. une gestuelle de mouvement vers la gauche ou vers la droite et un contact [38] (Fig. 6.3).
3. une gestuelle de mouvement vers la gauche ou vers la droite et un contact avec les doigts écartés [38] (Fig. 6.4).
4. une gestuelle de mouvement vers la gauche ou vers la droite et un contact avec les doigts en V [38] (Fig. 6.5).
5. déplacement par translation main de gauche à droite ou de haut en bas [55].
6. retournement de la main [55] (Fig. 6.6).
7. action d'ouvrir et de fermer la main [55] (Fig. 6.7).
8. déplacement de la main de haut en bas en pinçant le pouce et l'index [55] (Fig. 6.8).

Notons que certaines images présentées n'offrent pas forcément le même point de vue. L'idée est de rendre l'image plus facilement reconnaissable. Mais dans le véhicule, le choix du sens de la main au moment du geste doit être pensé pour que le geste puisse se faire naturellement.

Il existe d'autres gestuelles possibles mais nous nous limiterons à celles-ci.





FIGURE 6.1 – Valeurs 1,2,3,4,5[38]



FIGURE 6.2 – Fist, flat, grab, pich, point[38]

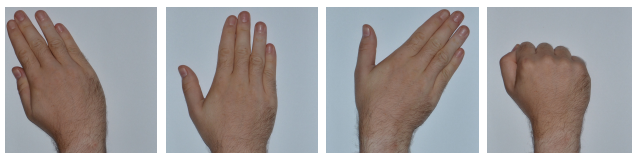


FIGURE 6.3 – Gauche Central Droite Sélection[50]



FIGURE 6.4 – Gauche Central Droite Sélection[50]

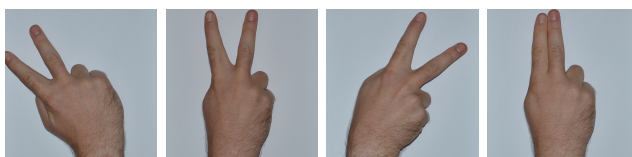


FIGURE 6.5 – Gauche Central Droite Sélection[50]



FIGURE 6.6 – Retournement de la main[55]



FIGURE 6.7 – Ouvrir, fermer et ouvrir la main[55]



FIGURE 6.8 – Déplacement de haut en bas[55]

On peut cependant remarquer plusieurs choses. S’il avait déjà été remarqué que le choix de la gestuelle choisie doit se faire en considération de la technologie utilisée, certains gestes bien qu’ayant une signification différente sont en réalité très proches voire totalement identiques. En effet, donner l’indication ”2” avec deux doigts levés [38] est équivalent au ’V’ proposé par [51]. D’autres gestes bien que différents sont assez proches et peuvent soit se révéler fort difficiles à différencier, soit donner lieu à une erreur de geste du conducteur. Enfin, pour ce qui est des mouvements de la main, nous pouvons également rappeler que l’état de la route ou les dispositifs ralentisseurs lorsque le conducteur effectue une action peut provoquer un mouvement involontaire qui engendrerait une action non souhaitée du véhicule. Si le document [22] nous met en garde sur les problèmes que les ralentisseurs provoquent sur les touchscreen, il est probable que le système de gestuelle en l’air sera tout autant voir plus impacté. Le document [54] va même plus loin en pointant l’impact sur la gestuelle des accélérations, freinages et changements de direction.

### Micro-gestuelle

La micro-gestuelle limite fortement la liste des gestes possibles mais permet au conducteur de garder ses deux mains sur le volant. Cet avantage n’est pas négligeable dans une situation de conduite difficile. Pour cette technique, la technologie utilisée pour détecter le geste doit être placée pour pouvoir détecter la position des mains sur le volant. De plus le fait que le conducteur doive tourner le volant dans un virage peut gêner la détection du geste.

Dans le document [27], on propose de lever ou de baisser un doigt de chaque

main pour effectuer une action.

D'autres gestuelles intermédiaires sont également possibles. Par exemple, le geste est effectué alors que la main ou le poignet est en contact avec le volant mais sans le tenir. Cette possibilité peut avoir son intérêt car la main utilisée pour commander reste très proche du volant ce qui permet au conducteur de reposer les deux mains sur le volant plus rapidement.

Une dernière technique possible que l'on peut également proposer [2] est de déplacer les mains du volant. Cela peut se faire avec des capteurs positionnés sur le volant : faire glisser une main, ou émettre une pression à un endroit précis du volant permet d'exécuter une action. Elle nécessite peu de technologie et sera moins sujette à des problèmes visuels. Le désavantage de la technique est le changement de position du volant lorsqu'on le tourne.

## 6.5 Reconnaissance vocale

L'utilisation de la reconnaissance vocale pour communiquer avec le véhicule est une technique très intéressante et a fait l'objet de beaucoup d'études. Il faut dire que les assistants vocaux sont devenus très populaires et disponibles dans de nombreux environnements. Le principal problème est le bruit présent dans l'habitacle qui rend la perception du son difficile pour le système de reconnaissance. Ce bruit peut provenir de la voiture sur la route, d'une fenêtre ouverte à grande vitesse, ou de la musique jouée dans le véhicule par le système d'infotainment. Cependant, l'utilisation d'un système de réduction de bruit permet de réduire la mauvaise compréhension par la voiture de 66%[49].

Ses avantages sont nombreux : le plus important est qu'elle permet de diminuer fortement la perte de vision de la route ou l'obligation d'enlever les mains du volant. A noter que cela peut quand même induire des interactions visuelles-manuelles pouvant être distrayantes[24]. Comparé à l'écriture sur un touchpad, la reconnaissance vocale se montre beaucoup plus sécurisante grâce à sa faible demande visuelle[13]. Il y a cependant toujours des recherches en cours pour identifier la charge mentale qu'elle peut provoquer sur la concentration du conducteur[42, 24, 41]. Même à commande vocale, une tâche peut se révéler trop complexe et avoir une mauvaise influence sur la conduite[24].

Un autre problème est que pour éviter toute action intempestive, on propose l'utilisation d'un bouton de type PTT (Push To Talk) [36]. Mais dans le cas de l'utilisation de ce système en chatbot, cela rend la communication moins naturelle.

Bien utilisée, par sa facilité et sa rapidité d'utilisation, la commande vocale permet de nombreuses applications comme par exemple, identifier le conducteur.[49] On peut également initier une communication proactive avec le conducteur, mais il ne faut pas perdre de vue que cette conversation est la seconde tâche. La conduite doit rester la première tâche. Cette technique peut permettre de réduire les accidents mais pas les supprimer [41].

### 6.5.1 Techniques de reconnaissance vocale utilisées

A cause des problèmes de bruit dans l'habitacle du véhicule, il est nécessaire d'avoir un système particulièrement efficace pour la reconnaissance vocale. Un ASR peut être utilisé mais il faut un moyen de diminuer le bruit ambiant. Depuis les années 70, on essaye avec différents algorithmes de diminuer le bruit et d'isoler la parole[32].

En ce qui concerne la reconnaissance de la parole, on utilise le plus souvent des techniques de machine learning.

## 6.6 Regard

Le système de commande par le regard propose une logique assez simple. Une caméra détecte les mouvements des yeux du conducteur pour en déduire ce qu'il regarde. S'il regarde un objet activable comme un bouton sur la console centrale du tableau de bord, il peut l'activer. Un problème se présente car ce n'est pas parce qu'on regarde la console centrale qu'on souhaite sélectionner quelque chose. Il faut donc une autre technologie pour permettre l'activation. Cela démontre que cette technologie ne se suffit pas à elle-même.

Cependant, cette technologie peut se révéler assez précise avec une détection d'angle de vision avec une erreur de  $0.4^\circ$ [3]. Cependant, cette valeur augmente fortement dans le cas où la position du conducteur vient à changer[3]. Une solution pour pallier ce problème peut être d'augmenter la taille des boutons que l'on peut sélectionner avec un tel système. Cependant, des algorithmes de correction peuvent diminuer le nombre d'erreurs de sélection de 50%[3].

Le réel avantage d'une telle technologie est sa vitesse d'action. Un regard simple prend 300ms[3], ce qui est très rapide[36]. L'intérêt dépend surtout du type de tâche que l'on souhaite effectuer[36]. Son principal défaut est qu'elle requiert de quitter la route des yeux, ce qui est dommageable pour

la conduite. De plus, les tests ont montré que les utilisateurs étaient assez réticents à utiliser ce système[40, 29].

## 6.7 Roulette

L'utilisation d'une roulette positionnée par exemple près du levier de vitesse présente l'avantage de pouvoir effectuer une opération de scrolling rapidement. Ce système est proposé dans des comparatifs mais bien que certains utilisateurs l'aient apprécié, les performances (vitesse, charge mentale, ...) se sont révélées moins bonnes face aux autres interfaces testées.

Il faut cependant reconnaître qu'il peut se révéler intéressant pour la sécurité car il est possible de le manipuler en gardant les yeux sur la route.

## 6.8 Synthèse

Les interfaces de communication vers les véhicules sont assez nombreuses. Chacune a ses qualités et ses défauts et si certaines comme les touchscreen sont depuis de nombreuses années utilisées dans nos véhicules, d'autres en sont encore au stade expérimental. D'autres enfin comme les boutons tendent de plus en plus à disparaître au profit de systèmes plus évolués permettant soit de simplifier les opérations, soit de réduire l'espace exploité. L'objectif général de ces interfaces vise la sécurité. Dans le cas spécifique des interfaces de communication du conducteur vers le véhicule, le but est de permettre au conducteur de communiquer en limitant l'impact sur plusieurs éléments :

1. Le regard de la route
2. Les mains sur le volant
3. La charge mentale

L'utilisation d'une technologie qui semble a priori trop lourde pourrait avec des modifications se révéler plus acceptable. On pourrait par exemple combiner ces interfaces pour améliorer leur efficacité. Par exemple l'utilisation d'un bouton PTT [36].

Enfin, les recherches ont montré que sur ces interfaces de communication du conducteur vers le véhicule, tous les participants aux tests n'ont pas montré les mêmes intérêts pour les interfaces. Même si cela semble être moins le cas d'un point de vue objectif, subjectivement, les préférences ne vont pas toujours vers les mêmes interfaces. Certaines interfaces, bien que pouvant se montrer utiles, sont rejetées par les utilisateurs[40]. Cela montre qu'il ne faut pas forcément se limiter à une interface, mais qu'il peut être intéressant d'en proposer plusieurs, pour peu qu'elles n'entrent pas en conflit. Le conducteur pourrait alors choisir celle qui convient le mieux de son point de vue.

## Chapitre 7

# Les interfaces de sortie

### 7.1 L'écran - Touchscreen

Comme cité précédemment, beaucoup de véhicules sont dotés aujourd'hui d'un écran tactile touchscreen. Ils offrent parmi de nombreux avantages, celui d'avoir accès à un grand nombre d'informations dans le véhicule comme la gestion des plus de 700 contrôles environnementaux que les constructeurs proposent [47]. Ils sont très pratiques pour afficher tous les éléments nécessaires : boutons, clavier pour la saisie de texte, liste avec scrolling pour la musique, la gestion du GPS, ou la navigation dans les options du véhicule. La liste de ce qu'on peut faire est énorme. Non seulement ils s'adaptent assez bien à toutes les situations visuelles de jour ou de nuit, mais encore ils peuvent facilement s'adapter aux technologies d'entrée proposées précédemment.

Dans le cas de la technique de sélection en l'air sans toucher proposée précédemment, l'écran permet de gérer les différents points de sélection proposés par cette technique[1].

Dans le cas d'autres types de sélection plus classiques, les touchscreen peuvent s'adapter au besoin en augmentant la taille des boutons de contrôle affichés.

Mais du point de vue de la sécurité, ils posent un défi majeur. L'utilisation du regard pour effectuer une action a aussi ses inconvénients. Le procédé a pour défaut d'accaparer fortement l'attention visuelle du conducteur. Bon nombre d'accidents peuvent lui être directement imputés [18].

De plus, pour l'entrée de texte, les recherches ont montré que la distraction visuelle qu'il provoque ne permet pas une conduite sans risque pendant son utilisation, que ce soit via un clavier visuel, ou en proposant une zone d'écriture[13]. Enfin, l'utilisation du touchpad exige que l'utilisateur enlève

une main du volant pendant une période pouvant être trop longue.

## 7.2 Système à vibration

Ce système fonctionne en positionnant un système vibrant sur le volant permettant au conducteur de recevoir une information lors de la vibration. Cette interface peut être très utile : le conducteur peut recevoir l'information de la validation d'une opération demandée sans influencer sur sa vision [55] et sans le forcer à retirer les mains du volant.

Elle ne permet cependant pas de faire beaucoup plus que valider une action. De plus, elle a pour défaut de pouvoir être mal perçue si les vibrations de la route sont trop importantes[46, 6].

L'idée de mettre plusieurs système à vibration sur le volant a été proposée [44] mais le retour a montré qu'il est difficile de savoir quel système a vibré.

## 7.3 Utilisation de solénoïdes

Une technique un peu spéciale proposée dans le document [44] est l'utilisation d'un système de feedback cutané par plusieurs solénoïdes. Cette technique se base sur le fait que des recherches ont démontré qu'un système de feedback par vibration est peu efficace : si on émet des vibrations sur différents endroits du volant, il est très difficile pour le conducteur de percevoir l'endroit précis où la vibration a été actionnée et donc de comprendre le message que l'on souhaite transmettre[44]. Un système de plusieurs solénoïdes est donc positionné à différents endroits du volant en fonction des endroits de la main qu'on souhaite toucher. Ce système offre l'avantage de permettre d'émettre plusieurs signaux tactiles facilement reconnaissables par le conducteur. Une limite à ce procédé est que le conducteur n'a pas toujours les mains au même endroit sur le volant. S'il doit tourner le volant dans un carrefour ou lors de manoeuvres, il pourrait ne pas percevoir une information transmise à cause de la mauvaise position de sa main ou ne pas pouvoir la comprendre. Dans le document [44], il est surtout employé comme comparatif avec d'autres technologies mais pourrait faire l'objet de recherches plus poussées.

## 7.4 Système ultrasonique

Les systèmes à ultrasons permettent d'envoyer une sensation de friction sur la main du conducteur pour lui transmettre une information[45]. L'action se fait via un quadrillage de haut-parleurs ultrasoniques transmettant une information. "Focused ultrasound haptics was invented in Japan in 2010

and commercialized in 2013 by Ultrahaptics (now Ultraleap) in the UK” [55].

Ce système permet de transmettre plus qu’une simple information de validation, car le quadrillage peut par exemple transmettre un signal circulaire dans le sens horlogique ou anti-horlogique. Cela peut, par exemple, indiquer la validation ou le rejet d’une commande.

Ce système a aussi l’avantage de ne pas provoquer d’interaction visuelle ou sonore et donc de présenter un impact très faible sur la conduite. Des études ont montré un intérêt grandissant pour cette technologie[45]. Elle se marie extrêmement bien avec les technologies sans contact et ne demandant pas d’attention visuelle.

L’avantage de cette technologie est qu’elle permet d’envoyer un feedback à des actions et donc de réduire la charge mentale sur le conducteur[45].

## 7.5 Système Sonore

Les systèmes sonores ne sont pas neufs dans les véhicules. On les utilise par exemple pour indiquer un problème au conducteur. Presque tous les véhicules aujourd’hui envoient un signal d’alerte sonore pour indiquer au conducteur qu’il a oublié sa ceinture de sécurité. Ils ont l’avantage de ne nécessiter que peu de technologie : le système radio et des haut-parleurs. Les signaux d’alerte sonores sont très pratiques car ils permettent d’attirer l’attention du conducteur alors qu’il est concentré sur sa conduite. Bien que ce système d’avertissement soit très efficace, il importe de ne l’utiliser d’en cas de réelle nécessité. Envoyer des signaux intempestifs dérangerait le conducteur, qui pourrait les ignorer ou gêner une conversation téléphonique ou avec les autres passagers. Bien choisir son alerte est également important. Un signal d’alerte mal choisi peut être mal interprété, voire ignoré[37].

Les systèmes sonores peuvent être de différents types : note, voix, ou spearcons(voir plus bas), icônes auditives[37]. Il importe de bien choisir le type de signal que l’on souhaite envoyer.

Les alertes et messages audio ont pour avantage de limiter la charge visuelle. Mais ce n’est pas un système parfait en matière de sécurité pour la conduite. Dans le document [24], on cherche à évaluer la charge mentale des messages audios. Les résultats indiquent que la charge mentale n’est pas négligeable et qu’il faut utiliser ces systèmes avec parcimonie pour ne pas altérer la qualité de conduite.



### 7.5.1 Spearcons

Les spearcons sont des modifications vocales de paroles(TTS) pour qu'elles soient lues plus vite. Les vitesses de lecture peuvent être variées. Elles ont l'avantage de permettre d'envoyer l'information plus vite. De plus, on remarque que le temps de réponse à ce système est également plus rapide en cas d'alerte[37]. Ils sont assez proches avec un temps de 40 et 60% de la longueur originale.

### 7.5.2 Sonification

On entend par Sonification le fait de transmettre un signal audio non vocal pour transmettre une information[5]. Le choix d'une ou plusieurs notes pour transmettre une information est assez courant mais il est possible d'aller plus loin. On peut par exemple rendre le son intermittent, augmenter ou diminuer le volume sonore. Tout comme un système intermittent, le fait d'augmenter ou de diminuer la fréquence de la note permet de transmettre une information.

Dans le document [5], on propose l'utilisation de ce système pour déplacer un objet virtuel en limitant la distraction visuelle. Même si le système fonctionne, le faible taux de réussite des tests a démontré une piètre précision.

### 7.5.3 Voix

Les systèmes sonores vocaux sont aujourd'hui très fréquemment utilisés pour les assistants vocaux. On utilise des techniques de type text to speech TTS. Ils permettent de produire un signal vocal parlé sur base d'un texte.

On les utilise souvent avec la reconnaissance vocale décrite précédemment. Ces deux techniques de reconnaissance vocale et de production de voix sont très fréquemment utilisées ensemble car elles permettent d'exploiter le même canal. La voix est prédominante dans les interactions entre personnes et c'est aussi l'un des modes de communication les plus efficaces. [41]. Elle permet de garder les mains sur le volant et les yeux sur la route. [36]. Mais la charge mentale exercée n'est pas négligeable.

Un autre souci est que l'environnement d'un véhicule est en général fort bruyant : bruits de moteurs, friction de l'air extérieur, air conditionné, chauffage, roulements. L'utilisation de systèmes vocaux peut devenir difficilement compréhensible par le conducteur[8]. Une solution peut être de modifier le son avant sa transmission dans les haut-parleurs pour rendre la voix plus intelligible dans cet environnement[8].

## Les chatbots

Les chatbots constituent une évolution mélangeant la reconnaissance vocale avec un outil de synthétisation vocale "text to speech" qui permet au conducteur de tenir avec le véhicule un dialogue simple ou complexe.

"Using Personal Assistants (PAs) via voice becomes increasingly popular and available in multiple environments"[41].

Dans les véhicules, les chatbots trouvent également leur place. Ils peuvent être utiles pour des raisons de sécurité, avertir le conducteur d'un danger par exemple, ou prévenir les accidents[41], mais aussi fournir des informations de navigation, donner les informations d'un message, la gestion du véhicule, etc. Ils offrent l'avantage de permettre une communication sans enlever les mains du volant pour peu que le système d'activation de la parole permette d'éviter une activation intempestive.

Cependant, comme cité précédemment, l'environnement sonore du véhicule peut rendre difficile la communication entre le chatbot et le conducteur. De plus, la charge mentale de la conversation sur le conducteur avec un chatbot peut être importante[41].

On en arrive donc à un certain conflit : d'une part, des recherches qui étudient la charge mentale sur les conducteurs reconnaissent l'intérêt sur la charge visuelle et le peu de mouvements qu'ils requièrent. Elles mettent pourtant en garde sur la charge mentale induite et le fait que cela peut provoquer des accidents[41, 16]. Un dialogue bien pensé est important[16]. De plus la charge mentale varie selon le type de discussion. Des mots familiers se montreront mentalement moins lourds[24].

D'autre part, des recherches ont pour objectif d'utiliser ces chatbots pour améliorer la sécurité. On peut par exemple citer "Traeddy"[19], un chatbot dans un ours en peluche qui détecte les zones de ralentissement dans le trafic routier et initie une conversation avec le conducteur dans le but de diminuer le stress et l'énervement provoqué. Une étude propose même l'idée d'utiliser un siège comme une personne pour définir un chatbot[43].

Une autre utilisation possible proposée par le document [4], c'est l'intégration d'un système de chatbot dans le véhicule pour mettre en place une gestion du stress, d'énervement et d'agressivité. Ce système aurait pour but de gérer tout le côté affectif du conducteur via une communication étudiée.

Le document [12] pose la question du cas particulier des assistants vocaux dans les véhicules et de la satisfaction qu'il peuvent apporter. Le document fait référence à une série des années 80 présentant un véhicule aux capacités

utopiques<sup>1</sup>.

Un autre élément important est la confiance accordée à un chatbot. Dans le document [16], on démontre qu'il faut pousser le conducteur à gérer lui-même le dialogue pour améliorer la confiance qu'il accorde au système[16].

Dans les voitures totalement autonomes, des système de chatbots pourraient voir leur rôle évoluer et devenir plus important.

## 7.6 Afficheur frontal HUD (Head Up Display)

Les afficheurs Head Up Display (HUD) sont des écrans transparents permettant l'affichage d'informations directement dans le champ de vision du conducteur. Cela va de l'affichage de la distance de sécurité, information de vitesse, etc... On considère qu'il faut 1,1 seconde au conducteur pour lire une information sur le tableau de bord et 0,6 sec si cette information est lue sur un HUD[15].

### 7.6.1 Techniques

Deux techniques existent. Soit le HUD utilise un écran transparent. Dans ce cas, la loi interdit de positionner un tel objet au dessus d'une ligne donnée. Un tel système sera donc limité en hauteur.

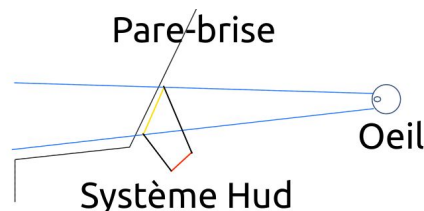


FIGURE 7.1 – Afficheur HUD

Ils ne vont généralement pas remplacer les composants comme les compteurs de vitesse du tableau de bord mais peuvent servir de complément pour améliorer la sécurité lors de l'accès aux informations[56].

Soit le pare-brise du véhicule sert de récepteur pour un projecteur. Dans ce cas de figure tout le pare-brise sert d'écran et permet donc l'affichage de beaucoup plus d'informations.

---

1. knight rider / K2000

### 7.6.2 Affichage d'informations sur le HUD

Dans le cas où le HUD utilise le pare-brise, on peut procéder de différentes façons pour afficher des données. On peut afficher des informations en 2D de façon classique. On peut aussi utiliser le HUD comme un système de réalité augmentée. Il est alors possible de s'en servir pour afficher des informations telles que la distance de sécurité avec la voiture qui précède. De plus un système de réalité augmentée exige que le HUD exploite une vision égocentrique.

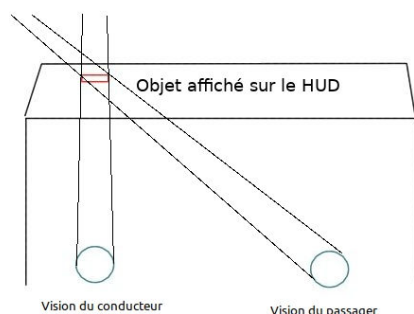


FIGURE 7.2 – Affichage égocentrique

Si le HUD propose une vision exocentrique, les passagers peuvent voir les informations comme le fait le conducteur. En affichage égocentrique (Fig. 7.2), les données sont affichées en fonction du point de vue du conducteur. Les informations affichées sont décalées si on se met du point de vue d'un passager [15].

Les HUD ont pour but initial d'afficher les informations utiles pendant la conduite. Le positionnement et la taille d'affichage de ces informations doivent être étudiés. Les informations importantes doivent se trouver au centre, en grande taille, mais suffisamment fines pour ne pas gêner la vue du conducteur. D'autres informations moins utiles peuvent être affichées dans un coin de la zone d'affichage [56].

Si on peut permettre d'afficher le titre de la musique écoutée dans le véhicule, on évitera d'afficher l'ensemble des paroles au milieu du HUD. Mal exploité, ce genre d'afficheur peut lui aussi représenter un réel danger. Trop d'informations ou des éléments en mouvement peuvent distraire le conducteur. De plus l'affichage d'informations peut rendre invisible au conducteur ce qui se trouve derrière [15]. Il est vital qu'un élément affiché sur l'écran ne puisse en

aucun cas dissimuler un élément extérieur important. L'utilisation du HUD en réalité augmentée pour afficher la distance nécessaire entre le véhicule et celui qui le précède peut poser ce problème. Voilà pourquoi les informations affichées sur un HUD doivent se limiter au minimum et viser la sécurité.

Des questions ont été posées à un panel de personnes pour savoir ce qu'ils voudraient voir affiché sur un tel système. Il en ressort que l'afficheur pourrait être utilisé pour une grande variété d'applications[9] : applications de sécurité, infotainment, monitoring du véhicule et la navigation. Ces informations seraient exploitées par les conducteurs mais aussi par les passagers.

L'utilisation du HUD pour de telles applications pose un gros problème de sécurité. Cependant, le document [56] propose une solution intéressante : proposer plusieurs configurations différentes du HUD en fonction de la situation. Une configuration sur route, une configuration sur autoroute qui prendrait en compte la vitesse et s'adapterait en conséquence, et une utilisation "parking" qui ne devrait plus gérer le fait que le conducteur conduit. Ce dernier point permettrait énormément de possibilités en réduisant les risques au maximum et pourrait même permettre d'afficher des jeux vidéos.

Enfin, il est probable que le HUD pourrait jouer un rôle beaucoup plus important dans le cas des véhicules autonomes.

## **7.7 Bandes de leds dans le champ visuel du conducteur**

Une autre technique possible est d'utiliser un affichage ambiant pour indiquer une information au conducteur. L'idée est d'attirer l'attention du conducteur sans vraiment le distraire en positionnant l'information à la périphérie de son champ visuel[21].

Ce système est moins distrayant qu'un touchscreen ou un HUD. Il n'a pas d'effet négatif sur la trajectoire et n'a pas les défauts d'un système acoustique[35].

Techniquement, l'idée générale est de placer une bande lumineuse de type LED dans la vision périphérique du conducteur. On peut la positionner à différents endroits du volant ou du tableau de bord. Le fait de positionner les leds sur le volant est moins demandeur d'attention car le conducteur aura moins tendance à regarder directement la bande lumineuse pour s'assurer qu'il ne s'est pas allumé.[21] Au-delà de cela, la position semble avoir peu d'importance.

Pour le choix des couleurs de la bande lumineuse, une couleur verte et blanche semblent être préférées par les testeurs selon le document[21]. Pour pouvoir transmettre plus d'information, plutôt que de changer de couleur, on peut proposer de transmettre un signal lumineux intermittent. Dans le document [21], ce système est utilisé par le système GPS du véhicule pour indiquer au conducteur la route à suivre.

Un tel système pourrait aussi trouver sa place dans les véhicules autonomes pour indiquer aux passagers le comportement prévu par le véhicule pour ne pas les surprendre, rendre la conduite plus agréable et augmenter leur confiance[39].

Reste le problème de la luminosité. Il semble que le meilleur moyen pourrait être de permettre au conducteur de pouvoir adapter sa configuration selon son ressenti.

## 7.8 Température

Une technique possible de diffusion d'information est la transmission de chaleur ou de froid. Le sens de la température est peu utilisé actuellement dans le véhicule sinon pour communiquer une sensation de bien être au conducteur. Le principe est le suivant : on chauffe le volant du côté où le conducteur doit tourner. Le "mauvais" côté est lui refroidi. On joue sur des paliers de 3° avec une zone neutre à 30°. Des études ont montré qu'un changement de direction via un changement de température peut être interprété correctement dans 94% des cas[6].

Si on l'utilise pour la sécurité, il a été démontré qu'une température chaude peut être associée à l'insécurité et au danger. Une augmentation forte de la température peut permettre d'avertir sur la sévérité du danger.

## 7.9 Transmission d'odeurs

Le principe est ici de propager dans l'habitacle du véhicule différentes odeurs correspondant à une information que le véhicule souhaite transmettre au conducteur[7]. On utilise un spray pour envoyer l'odeur souhaitée, citron, menthe poivrée, lavande ou rose[7]. Pour indiquer une vitesse excessive, on transmet une odeur de lavande dans le véhicule. Ce système a montré un impact positif sur le temps nécessaire au participant pour ralentir. Les conducteurs sont également restés moins longtemps en excès de vitesse.

Ce système présente l'avantage de ne pas exploiter les autres sens comme la vision et permet de garder les mains sur le volant.

## 7.10 Synthèse

Tout comme les interfaces de communication du conducteur vers le véhicule, les interfaces du véhicule vers le conducteur sont nombreuses et présentent chacune leurs avantages et leurs inconvénients. Elles sont utiles pour effectuer une action ou plus simplement pour valider une opération, elles sont sujettes aux mêmes recherches et aux mêmes objectifs de sécurité. Là aussi, on va chercher à diminuer leur impact sur la vision du conducteur, lui faire garder les mains sur le volant et diminuer la charge mentale.

L'idée d'utiliser plusieurs interfaces pour communiquer avec le conducteur permet de diminuer le risque qu'une information ne soit pas comprise à cause d'un environnement inadéquat. Par exemple, trop de bruit provoqué par une fenêtre ouverte empêche d'entendre une information. Mais trop d'informations via différents canaux risque de provoquer une surcharge mentale importante sur le conducteur. Cela aboutirait à un stress, un énervement et rendrait la situation dangereuse.

De plus il faut prendre garde à transmettre correctement les informations et à utiliser un canal avec une communication adaptée au message à faire passer en fonction de son importance et de l'urgence. Avertir d'un risque d'accident ne se fait pas de la même façon qu'on indique la réception d'un message. Si le choix du canal de communication est important, la façon de l'utiliser l'est tout autant.

## Chapitre 8

# Utilisation multimodale de ces interfaces

Après ce passage en revue des différentes interfaces, il apparaît que toutes ont leurs avantages et leurs défauts. Si chacune peut se montrer très utile dans certaines situations. Dans d'autres situations, elles peuvent se montrer inefficaces voire dangereuses.

C'est pour cela que bon nombre de recherches aujourd'hui tendent à utiliser ces interfaces dans un contexte multimodal dans le but de palier leurs défauts dans le cas où elles sont utilisées seules. Certaines interfaces ont même été conçues pour ne pas être utilisées seules mais pour fournir des informations déjà disponibles via d'autres interfaces. Par exemple, un HUD va permettre au conducteur d'avoir accès à des informations qui lui sont déjà fournies par d'autres composants plus difficiles d'accès. Le HUD sert donc à rassembler les informations utiles. Il ne remplace pas les composants traditionnels du tableau de bord mais montre ses avantages lorsqu'il leur est combiné[56].

Pour une même interface le constat est similaire. Différentes technologies de perception des gestes peuvent être exploitées simultanément pour pouvoir pallier une mauvaise perception d'information envoyée ou reçue par une technologie en utilisant une autre. L'utilisation d'une caméra combinée à une caméra infra-rouge par exemple peut permettre de limiter les cas où le système sera inefficace. Mais cela oblige à proposer une gestuelle qui sera perceptible par toutes ces technologies.

### 8.1 Combinaison de modalités

Il ya trois raisons à l'utilisation combinée de différentes modalités[31].



La première est que ce choix peut sembler naturel. Par exemple, si l'utilisateur transmet un ordre vocal au véhicule, il semblera pour lui naturel de recevoir sa réponse par un canal audio via une technique de synthèse vocale. C'est la technique de chatbot présentée précédemment.

La seconde est le choix. L'utilisateur a à sa disposition plusieurs interfaces d'entrées dans le véhicule lui permettant de transmettre une demande ou une information au véhicule. Il va donc choisir l'interface de communication la plus adéquate en fonction de l'information à transmettre, de la situation et de ses préférences.

La troisième est l'adaptation. Si une communication habituelle se fait par une discussion vocale entre le véhicule et le conducteur, on peut se retrouver dans une situation où la pollution sonore ne permet plus de tenir une conversation. A ce moment, une autre technique de communication va s'engager via par exemple le touchscreen du véhicule.

Au-delà de cela, le type d'information et l'importance de l'information transmise a aussi son importance dans le choix des interfaces de communication utilisées. Si l'importance de l'information transmise est faible, le choix doit être fait pour pouvoir transmettre l'information ou de la recevoir tout en permettant de réduire autant que possible de charge mentale du conducteur. Pour une information importante ou vitale comme une alerte pour la sécurité, on cherchera à transmettre l'information au plus vite et le plus efficacement possible.

C'est dans cette optique que beaucoup de recherches aujourd'hui posent non seulement la question de la nature des interfaces à utiliser entre elles, mais aussi de l'opportunité d'utiliser une seule ou plusieurs interface(s) d'envoi et ou de retour[27]. On peut par exemple allier la gestuelle, le regard et la parole pour commander le véhicule. Si certaines interactions peuvent sembler naturelles comme l'utilisation de la parole comme canal de communication par le conducteur, d'autres semblent plus surprenantes mais l'objectif est de trouver la combinaison d'interfaces qui permettra une communication limitant l'impact sur les sens du conducteur et sur sa charge mentale afin de ne pas le distraire.

Cette recherche sur l'utilisation combinée d'interfaces ne se limite pas à l'étude de l'utilisation d'interfaces d'entrée ou de sortie mais étudie surtout le choix de l'interface d'entrée utilisée pour telle ou telle interface de sortie ainsi que le montrent les documents qui suivent. Ces documents en arrivent souvent à proposer une utilisation conjointe de plusieurs interfaces d'entrée ou de sortie[27, 35, 44, 46, 37, 7, 21, 1].

Si les résultats de l'utilisation d'interfaces multimodales montrent une efficacité du système, d'autres interpellent sur le fait que lors des tests, les conducteurs n'ont pas tous émis les mêmes préférences. Cela prouve que certaines de ces interfaces peuvent s'avérer mentalement dérangeantes pour certains et utiles pour d'autres. On pourrait croire dès lors qu'il suffit simplement de permettre au conducteur de choisir. Mais cela nous ramène au problème de la configuration du véhicule avant de démarrer et au fait que, même si certains conducteurs n'ont pas aimé certains systèmes, les mesures objectives de ces tests ont montré leur efficacité.

Si le fait d'utiliser plusieurs interfaces est utile et peut améliorer l'efficacité ou la sécurité de la communication avec le véhicule, le fait qu'on l'utilise dans une communication multimodale a son importance. Il peut être important de ne pas utiliser de la même façon une interface. Par exemple, forcer le volume sonore pour le feedback d'une action gênerait le conducteur. Si la pollution sonore à l'intérieur du véhicule ne lui permet pas de percevoir l'information, il peut l'obtenir via un autre canal de communication. Sur ce point le document [35] propose l'utilisation d'une lumière ambiante par une bande led. Il conclut que même si le temps de réaction est plus important, les conducteurs préfèrent allier le feedback visuel par une lumière ambiante avec un système sonore. De son côté, le document [44] propose l'utilisation d'un feedback sensitif via un système à pression.

Un dernier élément à prendre en compte est la transmission ou la réception d'une information via des interfaces combinées. Le bouton PTT et le système LTT cité au paragraphe précédent en sont un exemple. Le bouton seul ou la parole seule ne sert à rien mais c'est la combinaison des deux éléments qui permet d'effectuer l'action.

## 8.2 Interfaces multimodales en entrée

Comme cela a été vu, chaque interface a ses qualités et ses inconvénients. L'habitacle d'un véhicule est soumis à des interférences provoquées par le véhicule ou son environnement qu'elles soient sonores, visuelles ou des vibrations. Aucun système de communication n'est épargné. Pourtant, certaines informations doivent être transmises le plus rapidement et le plus efficacement possible sous peine d'accident. De plus, même si ces interférences ne provoquent pas directement d'accident, elles peuvent engendrer de la frustration et une certaine déconcentration pouvant mener indirectement le conducteur à délaisser sa tâche première, la conduite, pour vérifier qu'une action a été effectuée ou pour la retransmettre. C'est pour cela qu'on va chercher à permettre au conducteur d'utiliser le moyen de communication qu'il souhaite ou qu'il trouve le plus judicieux compte tenu de la situation.

Mais toutes les interfaces ne vont pas avec tout et ne s'utilisent pas n'importe comment ni dans n'importe quelle situation. Ouvrir la fenêtre à un niveau voulu est peu pratique avec une commande vocale[27]. Le document [13] montre qu'un système de reconnaissance vocale s'avérera plus sûr pour la conduite qu'un touchscreen. Cependant il faut que l'environnement sonore du véhicule le permette. Dès lors, le conducteur doit pouvoir utiliser le touchscreen si besoin est.

### 8.3 Interfaces multimodales en sortie

Pour la transmission d'informations du véhicule vers le conducteur, on peut distinguer deux types de communication : la transmission d'une information et la transmission d'un feedback indiquant la bonne réception d'une commande. Au-delà de cela, l'importance et l'urgence de l'information doivent être prises en compte dans le choix des interfaces de transmission d'information.

Le document [21] avance l'idée d'utiliser des lumières dans le champ visuel périphérique du conducteur pour lui indiquer la direction à suivre communiquée par son système GPS. Le conducteur connaît la direction sans avoir à quitter la route des yeux et reçoit juste l'information nécessaire.

Le choix de l'interface est primordial pour faire passer l'information au mieux. Pour forcer un conducteur à ralentir, le document [7] propose d'utiliser conjointement une transmission d'odeurs, un message visuel "ralentissez" et un beep sonore. La transmission d'odeurs a pour avantage de transmettre au conducteur une information continue qu'il perçoit via un sens qui n'est pas exploité par les autres interfaces du véhicule. L'odeur transmise peut être mal interprétée par le conducteur et ne peut donc être utilisée seule. Il est également impossible de l'exploiter pour une situation d'urgence, mais seulement pour attirer l'attention sur un problème constant comme rouler à une vitesse excessive.

Cette logique de multimodalité a montré de très bons résultats. Les documents proposant plusieurs canaux de transmission d'informations confirment l'intérêt de cette technique[35, 7, 46]. Le conducteur perçoit l'information auditive alors que son attention visuelle est requise ailleurs. De plus cela permet de gérer les problèmes de pollution sonore, visuelle, ou autre du véhicule. Il est rare que tous les sens des conducteurs soient tellement perturbés par l'environnement que ce dernier ne puisse plus rien percevoir. (Dans une telle situation, l'important n'est pas que le conducteur reçoive certaines informations ou puisse les transmettre, mais bien de lui permettre par tous les

moyens possibles de rester maître de son véhicule). On peut, par exemple, renvoyer un feedback visuel et auditif.

### 8.3.1 feedback

Certaines recherches se concentrent sur le type de feedback à envoyer par certaines interfaces[44, 35, 28]. Cela montre bien que la relation entre le canal d'entrée du véhicule et le feedback renvoyé est critique. Dans le document [28], on indique qu'un feedback visuel et auditif a été préféré par les utilisateurs.

Les interfaces utilisées par le véhicule pour communiquer avec le conducteur, conçues exclusivement pour transmettre un feedback, ne peuvent généralement transmettre que peu d'informations. Les systèmes tactiles à vibrations ou à pression sur le volant figurent parmi les meilleurs exemples. Les différents tests effectués avec ces technologies font ressortir un point important : lorsqu'une information est transmise par le conducteur à un véhicule, un feedback du véhicule est nécessaire pour avertir le conducteur du succès de l'opération. Il est prouvé que la distraction provoquée est moindre si un feedback est transmis au conducteur[28]. De plus, les personnes ayant testé le système ont indiqué qu'elles préféreraient une interaction avec feedback plutôt que sans.

En conséquence, pour le retour, le véhicule va proposer par exemple, un retour sonore pour une action demandée, afficher un message sur un HUD et sur un écran du tableau de bord. Les différentes interfaces permettent au conducteur d'avoir le retour désiré, mais à condition que l'utilisateur ne soit pas envahi d'informations qui le déconcentreraient et dégraderaient sa conduite.

Dans le cas de l'utilisation de la gestuelle dans un véhicule, on a également besoin d'un feedback mais pas n'importe lequel[27]. Cette technique a pour but de limiter au maximum l'impact visuel, on aura donc tendance à éviter de renvoyer un feedback visuel direct qui oblige le conducteur à quitter la route des yeux. Mais le document [46] nous dit qu'un feedback non visuel a une charge mentale plus importante. Dans le document[27], la question est posée. On teste une information sur la console centrale. Le système peut émettre un feedback audio, des sensations sur la main et allumer une bande lumineuse. Il en ressort que les meilleurs résultats ont été obtenus par les interfaces dont le feedback n'est pas visuel[27]. Le document attire aussi l'attention sur le fait que le choix du canal de communication doit se faire en fonction de l'information à transmettre. C'est principalement le cas si la transmission d'informations initiale a pour objectif de permettre au conducteur de garder les yeux sur la route. Au pire, l'utilisation d'un HUD pourrait

être envisagée grâce à sa possibilité de garder les yeux dans la même direction.

L'article [44] propose un système de feedback sensitif à vibrations sur le volant ainsi qu'un système de lumière ambiante intéressants mais qui présentent comme tous les autres leurs défauts. L'article défend son utilisation au motif qu'ils ne demandent pas au conducteur de quitter la route des yeux. Mais cela n'est pratique qu'à condition qu'il soit possible au conducteur de pouvoir transmettre au véhicule son message sans avoir pour cela à quitter la route des yeux et au mieux, de pouvoir le faire en gardant les mains sur le volant. L'autre problème est que, s'il ne remarque pas le feedback soit parce qu'il ne l'a pas vu, soit parce que l'information n'a pas été reçue par le véhicule, l'attention du conducteur sera impactée car il voudra comprendre pourquoi il n'a pas reçu son feedback. Le document [46] étudie lui aussi la question du feedback mais indique que le feedback auditif induit une importante demande mentale.

## Chapitre 9

# Problématisation

L'espace du conducteur dans le véhicule est une zone très limitée et très protégée. Les réglementations en vigueur dans certain pays interdisent beaucoup de choses. De plus les normes généralement admises pour la gestion de l'environnement du conducteur émettent aussi de grandes réserves sur ce qu'il est possible de faire pour positionner les interfaces dans les véhicules. Pourtant, nous sommes depuis de nombreuses années face à une augmentation importante du nombre de contrôles nécessaires dans le véhicule. L'infotainment en particulier est très demandeur d'espace de contrôle.

Pour limiter autant que possible l'espace exploité par l'infotainment, les constructeurs utilisent depuis assez longtemps des touchscreen. Ils permettent de naviguer dans différents menus pour effectuer les actions souhaitées. Mais aujourd'hui, les constructeurs et les chercheurs cherchent de plus en plus à réduire l'utilisation de ces touchscreen. La raison est qu'ils ont tous perçu le danger de cette technologie. Si un conducteur qui connaît bien son véhicule peut aller chercher un contrôle sans quitter la route des yeux, voir en limitant fortement le retrait des mains du volant grâce aux boutons de contrôle sur celui-ci, il lui est impossible de manipuler un touchscreen sans quitter la route des yeux ni en gardant les mains sur le volant. Dans ces conditions et sans que le conducteur en prenne conscience, la manipulation de l'infotainment devient alors l'activité principale du conducteur et la conduite la seconde activité. De plus, il est impossible pour le conducteur de percevoir un danger en approche car il ne regarde plus la route. Le résultat est que de nombreux accidents sont provoqués par la distraction d'au moins un conducteur qui manipulait le touchscreen de son véhicule.

C'est ce constat qui pousse les constructeurs et les chercheurs à trouver de nouvelles techniques de communication avec le véhicule avec pour objectif de permettre au conducteur de pouvoir effectuer un maximum d'actions sans avoir à quitter la route des yeux, de garder les mains sur le volant et

de minimiser la charge mentale pour que la conduite reste l'activité principale du conducteur. Dans un véhicule parfait, ces trois éléments (yeux sur la route, mains sur le volant, charge mentale) devraient avoir un poids réduit à zéro pour toute autre action que la conduite pour le conducteur. La part des yeux sur la route reste la plus importante car dans le cas où un danger surviendrait, avec les yeux rivés sur l'infotainment, le conducteur ne remarquerait même pas celui-ci et ne pourrait donc l'éviter. Au-delà de cela, le document [44] va plus loin en indiquant que les tâches primaires et secondaires ne doivent pas employer les mêmes ressources, sous peine de conflits.

Il faut donc trouver d'autres interfaces qui exploitent le moins possible ces trois éléments. Certains documents étudient la gestuelle via différents capteurs. Leurs technologies peuvent varier : radar, caméra, vision infra rouge. chacune a montré ses avantages et ses faiblesses. Si certaines peuvent se montrer efficaces dans des milieux complexes tels que ceux d'un véhicule, elles se montrent souvent moins efficaces pour reconnaître les gestes et il est donc nécessaire de prévoir la gestuelle en conséquence. D'autres peuvent se montrer plus sensibles dans les milieux complexes mais permettre une reconnaissance plus précise de la gestuelle et permettent de reconnaître une gamme de gestes beaucoup plus large.

Si les technologies existantes ont leurs défauts, elles offrent aussi leurs avantages. Le plus grand est le fait que le conducteur peut effectuer un nombre assez important d'actions sans quitter la route des yeux. Cela permet de limiter voire de supprimer les risques qui en résultent que le conducteur quitte la route des yeux. En conséquence, cela augmente fortement la sécurité pendant que le conducteur effectue les actions souhaitées.

Les recherches ont démontré l'intérêt de ces technologies. Elles permettent une large gamme de gestes et donc, d'actions différentes. Cela est très pratique au vu du nombre de contrôles possibles dans un véhicule et qui doivent le plus aisément possible être accessibles au conducteur. Les constructeurs l'ont bien compris et étudient ces technologies ou les ont déjà intégrées dans leur gamme de véhicules. Elles permettent une large gamme de gestes sur une seule interface. Bien que généralement limité à la gestion de l'infotainment, il est reconnu qu'on peut s'en servir pour beaucoup d'actions. Même si ces technologies sont généralement limitées à la gestion de l'infotainment, elles permettent de limiter le nombre de boutons présents dans le véhicule, alors que l'espace est limité.

Une question qui se pose est le positionnement de ces interfaces. Tout comme les autres systèmes (boutons, touchscreen, ...), on ne peut positionner ces interfaces de reconnaissance n'importe où dans le véhicule. Les capteurs de gestes sont plus encore limités car ils doivent être placés de façon à permettre

au conducteur d'effectuer une gestuelle dans une zone facilement accessible. Le positionnement de cette zone doit être choisi en fonction de la main qui va devoir effectuer l'action. L'action dans les documents étudiés ne peut donc être faite que d'une main car la position du capteur empêche le conducteur d'utiliser l'autre main. De plus, Le conducteur doit toujours avoir une main sur le volant.

L'idée proposée ici est donc de savoir s'il est possible de permettre au conducteur d'utiliser ses deux mains pour effectuer une gestuelle. On ne parle pas ici de permettre d'effectuer une action avec les deux mains en l'air, car cela obligerait le conducteur à enlever ses deux mains du volant mais de permettre au conducteur d'effectuer une action de gestuelle via sa main droite ou sa main gauche. La question ici est de se demander s'il serait utile et pratique de permettre au conducteur d'effectuer des action de gestuelle de la main droite ou de la main gauche. Une telle idée ouvre de nouvelles portes. Quels seraient les avantages et des défauts d'une telle utilisation de cette technique ?

Premièrement le fait de permettre une gestuelle des deux mains alternativement en gardant toujours une main sur le volant peut-il se faire via un seul capteur. Si c'est le cas, où va-t-on pouvoir le positionner pour permettre au conducteur de pouvoir effectuer ses actions avec les deux mains. Si le capteur est positionné trop à droite ou trop à gauche, le conducteur ne pourra pas utiliser la gestuelle aisément, voire pas du tout. De plus, il faut que le capteur et la partie logicielle permettent de différencier les deux mains avec précision. Dans le cas où on choisirait de positionner deux capteurs, cela nécessiterait de leur réserver deux espaces distincts de part et d'autre du conducteur. Cette approche offre l'avantage de permettre une utilisation beaucoup plus aisée par le conducteur. On pourrait alors imaginer qu'un capteur ne pourrait recevoir de commande que d'une main et non de l'autre. Plus besoin dans ces conditions de les différencier, identifier le capteur d'origine revient à identifier la main qui a effectué le geste. Mais on doit alors se demander où devraient se positionner dans le véhicule les deux capteurs et quelle serait la zone d'utilisation dans laquelle la main va devoir effectuer son geste.

Pour ce qui est de la gestuelle, on se retrouverait avec plusieurs possibilités. La première consisterait à proposer les mêmes gestuelles pour les deux mains pour effectuer les mêmes actions. Cela permettrait au conducteur d'utiliser la main qu'il souhaite. Une deuxième possibilité serait de définir les mêmes gestuelles mais pour des actions différentes selon que le conducteur utilise sa main gauche ou sa main droite. La aussi cela pose la question de quelle gestuelle et de quelles actions pour quelles mains ? quels sont les avantages et les inconvénients de cette logique ? Une troisième possibilité



serait de fusionner les deux : on va permettre certaines action ou certains gestes d'être effectué le la main droite ou de la main gauche au choix, mais d'autres gestes n'auraient pas le même effet en fonction de la main gauche ou droite.

Au niveau des gestes, existe-t-il des gestes qui doivent se limiter à la main droite ou à la main gauche ? Y a-t-il des différences de perception du capteur utilisé en fonction de la main ?

Dans la plupart des documents étudiés, l'aspect de la sécurité est un élément central. Il est impossible de présenter une telle idée sans se poser la question de l'impact d'une telle évolution sur la sécurité. On tente d'évaluer l'impact sur la sécurité que pourraient avoir ces deux capteurs dans le véhicule. Le fait de pouvoir effectuer des commandes par gestes des deux mains peut-il avoir un impact sur l'attention du conducteur ? L'impact sur la sécurité est-il plus important que si on se limite à un capteur ne reconnaissant que la main droite ou la main gauche ?

Enfin se pose la question du type d'actions à effectuer ? Doit-on réserver certaines action à la main droite et d'autres à la main gauche ? Le fait d'avoir des gestuelles sur les deux mains ouvre-t-il la voie à l'utilisation plus étendue des actions à effectuer ? Est-il possible d'utiliser les actions d'une main pour un système dans le véhicule et une autre main pour un autre système ? Certains systèmes doivent-ils être limités à une main en particulier ? Par exemple, l'utilisation de l'infotainment doit-il être limité à une main et si oui, laquelle ? Enfin, doit-on permettre de gérer certaines interfaces avec les deux mains sachant qu'il est vital d'empêcher le conducteur d'enlever les deux mains du volant ? Comment bien choisir les actions possibles pour chaque main pour minimiser l'impact sur la sécurité ?

Un autre problème est que le volant dans les véhicules n'est pas situé du même côté en fonction de la conduite à gauche ou à droite. Quel est l'impact sur la reconnaissance de la gestuelle et le développement ? Le logiciel de l'interface/des interfaces doit-il être pensé en conséquence ?

Sur base de cette problématisation qu'est l'utilisation des deux mains pour l'envoi de commandes via une gestuelle en l'air, on en arrive à un nombre assez important de questions que nous venons de poser et auxquelles ce document va chercher à répondre au mieux dans la partie qui suit. Pour ce faire, nous proposerons le développement d'une programme de reconnaissance de gestuelle pouvant différencier une gestuelle main gauche et main droite.

## Chapitre 10

# Etude d'un système de reconnaissance de gestuelle pour voiture

Afin de tester la possibilité d'utiliser une gestuelle avec les deux mains dans un véhicule, il est proposé ici de développer un programme pour expérimenter cela. Le programme aura pour but de permettre d'effectuer des gestes des deux mains, il devra être capable de reconnaître la gestuelle proposée et d'indiquer l'action à faire en fonction de la gestuelle qui serait une gestuelle en l'air. Le conducteur devra enlever une main du volant pour effectuer une action mais doit pouvoir le faire sans quitter la route des yeux. Il est clair que le développement d'une telle technologie prend beaucoup de temps. Les recherches nécessaires pour le choix de la technologie de reconnaissance, son intégration dans un véhicule, les tests nécessaires de validation en laboratoire et les tests en conditions réelles prennent du temps et sont extrêmement coûteux.

Par conséquent, on se rapprochera fortement des technologies qui y sont utilisées dans les phases de simulation à ceci près qu'on va chercher à s'adapter au fait que l'élément étudié ici est l'utilisation des deux mains dans la gestuelle et bien sûr de l'impact que cela représente.

La première question est de savoir si on se dirige vers une gestuelle unique pour les deux mains. L'avantage de cette façon de procéder est qu'elle permet au conducteur de choisir la main qu'il trouve la plus adaptée pour exécuter le geste souhaité. Sachant qu'il existe 10% de gauchers, cela signifie que 10% pourraient avoir une préférence de choix de main différente du reste des conducteurs. Mais choisir cette option limite l'intérêt de la gestuelle à deux mains et il semble à ce stade plus intéressant de proposer une gestuelle différente en fonction de la main. Nous allons donc étudier une gestuelle de

commande et des choix de commande différents pour chaque main. Les tests sur l'application devront déterminer si cette idée est prometteuse ou non.

Il ne faut pas perdre de vue que les capteurs gauche et droit ne sont pas positionnés au même endroit dans toutes les voitures car il faut prendre en compte le fait que le volant est positionné à gauche sur les véhicules circulant à droite et à droite dans les véhicules circulant à gauche. Au niveau des interfaces, c'est une élément à prendre en compte. On parlera donc d'interface côté porte et d'interface au centre.

Dans un véhicule, l'idéal serait de positionner deux interfaces à portée directe de la main du conducteur. Du côté porte, on positionnerait le capteur au niveau des contrôles des vitres. La position naturelle de la main se positionnant aisément à cet endroit. Pour le positionnement du capteur pour la main qui est du côté central de la voiture, on pourrait le positionner soit au niveau du tableau de bord, au devant du levier de vitesse entre les sièges. Ce positionnement ne serait pas différent du positionnement voulu par les constructeurs utilisant une interface unique.

## **10.1 Actions proposées par le système de reconnaissance**

Dans les documents étudiés, il est peu fréquent que les système de gestuelle soient utilisés pour la conduite, en général, on va chercher à se cantonner à des éléments non critiques du véhicule tels que :

- L'ouverture ou la fermeture des vitres.
- Le système chauffage.
- Le système d'air conditionné.
- La gestion de l'infotainment.

Dans le document [55], on se pose la question des actions effectuées par le conducteur autres que conduire, soit contrôler la musique, sélectionner la température, répondre à des appels. Ces éléments peuvent être utilisés pour définir les choix des opérations possibles pour le contrôle de ces interfaces.

Partant du principe qu'on aurait une commande de chaque côté, il semble évident qu'on va chercher à sélectionner des actions spécifiques à la main côté porte et d'autres actions spécifiques à la main côté centre. Pour ce qui est de la main côté porte, une idée intéressante serait de permettre d'exploiter entre autres les commandes de gestion des vitres. Le conducteur étant habitué à contrôler les vitres de cette main, il ne sera pas gêné à l'idée d'utiliser ce système pour peu que la gestuelle soit simple.

Partant sur la même logique, l'autre main devrait être utilisée pour la gestion

de l'environnement chauffage/air-conditionné, ou de l'infotainment. Cependant, dans la majorité des véhicules la gestion de la température de l'environnement est une opération moins dangereuse et moins complexe que la gestion de l'infotainment. Il est donc évident que la gestuelle proposée pour ce capteur doit se concentrer sur des action pour l'infotainment. On pourrait alors se concentrer sur les action souvent citées vues précédemment : augmenter ou diminuer le volume, répondre ou raccrocher le téléphone, choisir le canal radio, musique suivante, musique précédente... Le choix de contrôler le chauffage et/ou l'air conditionné serait donc accessoire voire tout simplement ignoré.

## 10.2 Gestuelle choisie

La gestuelle de contrôle doit être choisie avec soin. Comme cité précédemment, il est préférable de réduire la liste des gestes possibles au minimum, car la gestuelle employée demande un travail de mémorisation des commandes. Trop de commandes et donc trop d'actions risque de rendre le système trop complexe à utiliser. Cela peut provoquer des erreurs de gestes ou l'impossibilité pour le conducteur de se souvenir du geste souhaité. Le choix de la gestuelle est donc primordial[2].

On va donc se concentrer sur une gestuelle la plus simple possible. De plus cette gestuelle doit également être gérée en fonction de l'interface utilisée. L'interface utilisée qui, nous le verrons plus tard, ne devrait pas poser de problème pour peu que la gestuelle ne soit pas trop complexe. On retrouve ici deux bonnes raisons de proposer une gestuelle simple : faciliter la perception par l'interface et faciliter la gestuelle effectuée par l'utilisateur. Cette gestuelle doit également être logique. Comme vu précédemment, certains gestes sont plus logiques que d'autres pour effectuer certaines actions. Si certains gestes se montrent plus logiques pour un "oui" et que d'autres se montrent plus logique pour un "non"[55], on peut raisonnablement penser qu'il en est de même pour d'autres actions. Dans le cas présent, proposer un geste qui demanderait de baisser la main pour lever les vitres semble fort peu logique.

Enfin, il faut que la gestuelle soit simple et facile à exécuter, une gestuelle demandant au conducteur un effort mental ou de se contorsionner serait dangereux. Certains gestes sont possibles mais trop complexes à effectuer mentalement. Un geste circulaire par exemple est déjà fortement déconseillé. D'autres gestes ont une demande mentale beaucoup trop forte. Par exemple, ce geste qu'on appelle souvent "le salut Vulkan". (Fig. 10.1)



FIGURE 10.1 – Geste trop difficile

Un autre problème est que, même si la technologie employée est assez précise, il convient de ne pas utiliser de gestuelles trop proches les unes des autres. Par exemple, un geste avec les doigts joints et les doigts écartés pourrait être mal interprétés par le système qui effectuerait une mauvaise action. (Fig. 10.2)



FIGURE 10.2 – Gestes trop proches

Un autre problème sur le geste proposé sur l'image (Fig. 10.2) est que la position des doigts est assez proche de la position naturelle. Cela signifie que si le conducteur passe sa main par inadvertance dans la zone de perception du capteur et que la position de celle-ci est trop proche d'une gestuelle, il va déclencher une opération non sollicitée.

Un dernier problème est d'éviter que le conducteur doive effectuer un geste trop contraignant. Par exemple, commander le son de l'infotainment avec le pouce levé ou baissé serait là aussi fort contraignant car cela demanderait une contorsion importante pour le pouce baissé.

On va donc chercher à se diriger vers une gestuelle simple, logique et aisée pour le conducteur. Pour la partie côté fenêtre, pour gérer l'ouverture des vitres, par exemple, on peut proposer de combiner un geste des doigts et de la main. La chose est déjà assez complexe mais devrait se révéler pour l'utilisateur assez logique. Par exemple, on considère que chaque fenêtre à un numéro :

1. Fenêtre côté conducteur.
2. Fenêtre côté passager.
3. Fenêtre arrière côté conducteur.
4. Fenêtre arrière côté passager.

Le document [38] propose une gestuelle pouvant être facilement adaptée à un panel d’actions assez large dans un véhicule et dont l’objectif est assez proche. Bien qu’ayant fait le choix de la gestuelle pour un système de reconnaissance différent et pour une analyse via du machine learning, on peut penser que la technologie et l’algorithmique que nous allons utiliser se montrerait tout aussi efficace. On peut donc reprendre tout ou une partie de cette gestuelle pour définir nos actions ou du moins se baser sur celles-ci.

On propose donc de choisir la vitre à commander en effectuant le geste avec les doigts et en levant ou en baissant la main pour lever ou baisser la vitre choisie. L’action proposée par ce geste serait continue. La personne baisserait ou lèverait la main et garderait le geste des doigts tant qu’il souhaite voir la vitre continuer à se lever ou se baisser. Ce choix, s’il permet d’effectuer une action continue, a le défaut d’empêcher le conducteur de remettre la main sur le volant tant que l’action est en cours. S’il veut ouvrir ou fermer une fenêtre complètement, il devra enlever la main du volant pendant une période de plusieurs secondes.

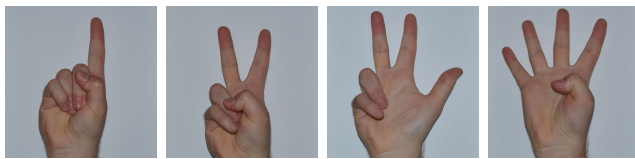


FIGURE 10.3 – Choix de vitre 1,2,3,4

Pour ce qui est de la main côté centre, on proposerait de gérer l’infotainment. On utiliserait la même gestuelle dans un premier temps pour permettre de choisir les action (Fig. 10.4). On peut proposer par exemple :

1. Radio.
2. GPS.
3. Téléphonie.
4. Musique memory stick.

On proposerait deux autre gestes de validation, à savoir : toucher pouce-index pour valider, et toucher pouce-majeur pour rejeter/annuler (Fig. 10.4). Pour le reste, on pourrait permettre la navigation dans les options de l’infotainment avec les gestes 1,2,3 et 4.

Cette logique se rapproche très fortement de la logique présentée dans le document [55]. Cependant, pour qu’une telle gestion soit utile, il faudrait proposer un système qui permette au conducteur de connaître les options proposées par l’interface d’infotainment sans devoir quitter la route des yeux pour ne pas perdre l’intérêt sécuritaire de cette technique.

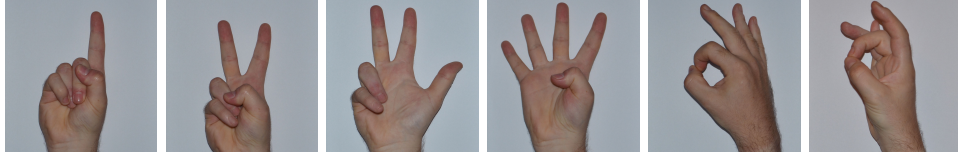


FIGURE 10.4 – gestuelle main côté centre du véhicule

### 10.3 Confirmation du geste

Comme cela a été vu, la charge mentale augmente si le conducteur ne reçoit pas de confirmation du geste effectué[28]. Il faut donc une technique pour confirmer la bonne marche de l'exécution. Pour éviter d'accaparer l'attention visuelle du conducteur on va proposer un feedback sonore. Il devra être différent en fonction de l'action effectuée. On pourrait choisir d'utiliser des spearcons qui ont l'avantage d'être assez rapides à jouer. Cependant, jouer un spearcon différent pour chaque action aurait pour effet que le conducteur ne pourrait certifier qu'il a effectué la bonne action. On obtiendrait alors les mêmes effets sur sa charge mentale que s'il n'avait reçu aucune confirmation.

On pourrait donc donner une confirmation vocale pour certaines actions. Par exemple, le choix du téléphone enverrait dans l'habitacle le son "téléphone". De cette façon, le conducteur serait certain d'avoir ou non sélectionné la bonne action.

Enfin, au sujet des vitres, on peut se demander s'il est vraiment nécessaire de confirmer la bonne marche de l'opération vu que l'ouverture ou la fermeture des vitres fourni au conducteur un feedback ambiant visuel ou sonore.

# Chapitre 11

## Implémentation

### 11.1 Choix du matériel et des logiciels

Le but de ce travail est d'étudier la possibilité de proposer une reconnaissance de gestes des deux mains et de permettre d'accomplir des actions différentes en fonction de la main qui a effectué le geste. Le problème ici est qu'on ne dispose pas d'un véhicule modifiable pour effectuer le test, ni des autorisations ou homologations nécessaires pour faire circuler un tel véhicule modifié sur la voie publique, ni d'un environnement de test de véhicule. Nous nous limiterons donc à une phase proche des études de nouvelles technologies en laboratoire. Il est clair que cela ne peut suffire à confirmer que cette technique soit viable mais peut au moins permettre de déterminer si elle mérite une étude plus poussée comme une simulation, voire un test en condition réelle. Le développement se fera donc avec le matériel suivant : une architecture X86 (64 bits) avec une plateforme JAVA 8 sous Linux Mint 19.2. On s'en servira pour recevoir les données du/des capteurs de reconnaissance.

L'interface de reconnaissance des gestes devra fournir une liste de coordonnées représentant les différents points des mains importants pour la reconnaissance. Même si l'idée de base vise l'utilisation de plusieurs capteurs, pour cette phase de test, on se limitera à un seul module de reconnaissance. Cela implique qu'il doive être capable de donner un moyen clair d'identifier la main droite et la main gauche.

### 11.2 Interface de reconnaissance utilisée

Pour la reconnaissance des mains et des gestes des mains, il a été choisi d'utiliser ici le module Leap Motion. C'est un petit module qui permet de capter les mouvements de la main et des doigts. Il permet ensuite via des bibliothèques de récupérer ces informations dans un logiciel de développement. Ce module va transmettre différentes informations[11].



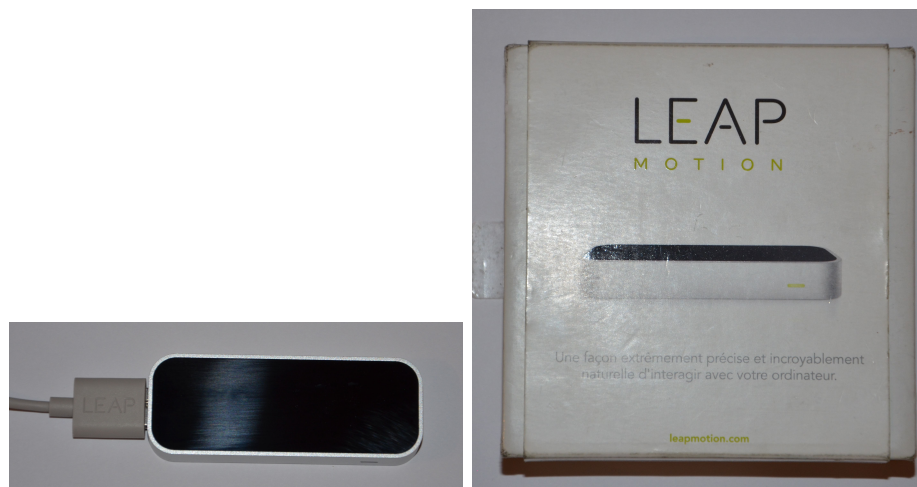


FIGURE 11.1 – Module Leap Motion

Le travail consistera donc à récupérer ces informations et à les décoder de façon à retirer les données nécessaires à la reconnaissance des différents éléments de la main. Une zone virtuelle est définie en face du capteur. C'est la zone de perception de la main. Si une main se place dans cette zone, elle est perçue par le module qui en retire les propriétés : position des doigts et position de la main.

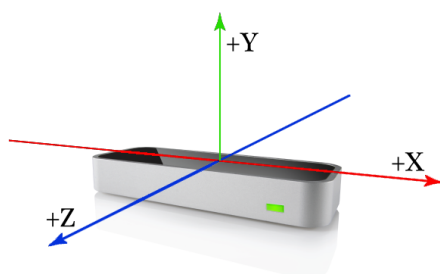


FIGURE 11.2 – Axes Leap Motion[11]

Tout d'abord, au niveau des différents doigts, le module va déterminer la position dans l'espace du début et de la fin des différentes phalanges qui composent la main ainsi que leur direction. Cette opération est répétée pour les différents os de chaque doigt. La position "fin" d'un os correspond au début de l'os suivant. Enfin, le module détermine la position et la direction du bout de chaque doigt.

Au-delà des doigts, le module détermine les différentes rotations de la main. Le module peut aussi reconnaître la main utilisée, main gauche ou main droite. Il est enfin capable de reconnaître les deux mains simultanément.

Cette possibilité de reconnaître les mains présentées devant le module sera très utile ici. S'il est possible de gérer plusieurs modules, on ne dispose que d'un module Leap Motion. Il faudra donc recevoir les données des deux mains avec le même module et définir s'il s'agit de la main gauche ou de la main droite pour agir en conséquence.

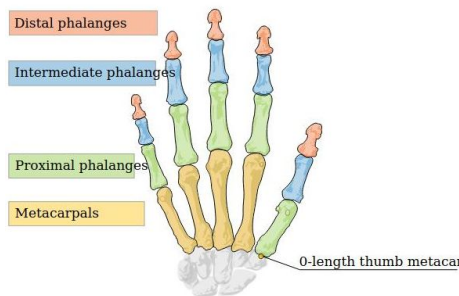


FIGURE 11.3 – Description d'un main

### 11.2.1 Performances du module Leap Motion

Pour effectuer des tests dans un milieu plus proche d'un laboratoire que d'une voiture, ce module convient très bien. Peu coûteux, il se montre assez performant et les données qu'il renvoie sont assez aisément exploitables par un logiciel pour différentes actions. Cependant, il est peu probable qu'un tel module puisse trouver tel quel sa place dans un véhicule. Si une telle technologie devait être utilisée dans un véhicule, vu les difficultés citées précédemment, il est probable qu'un autre système devrait être employé. Cependant, dans le contexte actuel, les informations transmises devraient permettre une reconnaissance assez poussée pour reconnaître les gestes cités précédemment. Le développement et les test devront cependant encore le confirmer.

## 11.3 Techniques utilisées pour la reconnaissance

Les informations reçues par le module sont transmises de façon assez brutes. Telles quelles, il est difficile de récupérer directement la position ou la direction des doigts. Un doigt positionné d'une façon précise sur la main ne donnera pas la même direction si la main est dans un sens différent. Si on fait le signe '1' du doigt (Fig. 10.4), la direction et la position des doigts renvoyées par le module seront différentes en fonction de la position de la main, elle-même. Il va donc falloir traiter les données reçues pour les rendre exploitables.

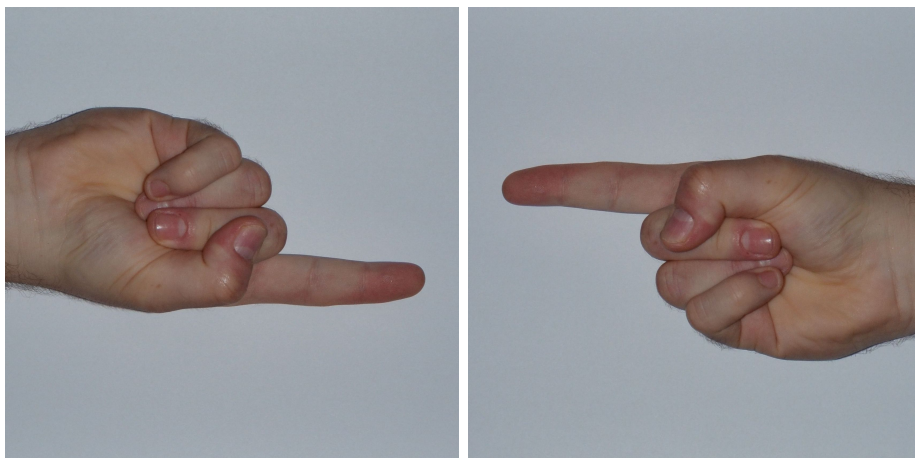


FIGURE 11.4 – image identique, geste identique, mais mains pointant dans deux directions opposées

Une technique possible présentée précédemment pourrait utiliser une reconnaissance par du machine learning[50, 38, 53], mais cela nécessiterait un entraînement conséquent pour être efficace[53]. Dans le cas présent, on cherche à reconnaître par des opérations mathématiques les gestes effectués. Pour ce faire, on définit un point central sur la main, le "palm position" : cet élément deviendra notre point 0,0,0.

A partir de là, on recalcule tous les éléments de la main. Avec les informations obtenues, on recherche ensuite les rotations de la main, pour effectuer une rotation inverse équivalente. De cette façon, la position de la main devant le capteur n'influence plus les valeurs perçues de la position des doigts.

On obtient donc un "package" de plusieurs éléments :

1. L'info de la main perçue main gauche ou droite
2. La position générale de la main dans l'espace
3. Les différentes rotations de la main dans l'espace
4. La position des différents points des os de la main par rapport au centre de la main
5. La position du bout de chaque doigt
6. La direction du bout de chaque doigt

Sur base de ces informations, on cherche à identifier un geste. Pour cela, on enregistre des caractéristiques spécifiques à chaque gestuelle. Par exemple, sur le geste "1" vu précédemment, l'index doit être pointé vers le haut, le majeur, l'annulaire et l'auriculaire doivent être pointés plutôt vers le bas.

Le pouce doit être dirigé vers l'intérieur de la main.

L'idée est donc de définir un range de valeurs acceptées pour certaines variables de la main, valeurs que nous possédons, les autres devant être ignorées. Chaque range se définit par trois valeurs, à savoir l'axe X, Y, et Z. Pour chaque valeur, on définit un minimum et un maximum accepté. Pour le geste "1" la direction Z de l'index doit avoir une valeur approximative de  $90^\circ$ , on va donc proposer un range de valeurs pour cette variable qui devrait obligatoirement être respecté. On peut donner un range de 70 à 100. Cette valeur est définie de façon arbitraire, les raisons et les conséquences de ce choix seront étudiées un peu plus loin dans ce document.

Pour ce qui est des autres doigts, il devront avoir une valeur inverse de l'ordre de  $-90^\circ$ . Partant du principe que le conducteur fera moins attention à les positionner correctement que pour l'index, on va leur permettre un range de valeurs plus large. Pour peu que l'angle soit négatif, c'est correct. Le range sera donc de  $-100^\circ$  à  $0^\circ$ . Pour ce qui est du pouce, c'est l'axe X qui doit être pris en compte. avec une évaluation similaire.

Le tout est de s'assurer de prendre suffisamment de valeurs pour éviter un faux positif ou plusieurs gestes différents reconnus, tout en étant pas trop restrictif pour permettre de reconnaître un geste aisément s'il se présente.

Pour la gestion des vitres qui a été choisie, il a été décidé qu'on baisserait ou lèverait la main pour commander l'action souhaitée. La position des doigts de la main étant connue et le geste reconnu, il ne reste qu'à identifier un mouvement de la main toute entière. Pendant ce mouvement, les doigts sont susceptibles de bouger mais ils doivent rester dans le range de valeurs qui permettent de reconnaître la gestuelle. L'extraction des valeurs étant faite, la position de la main et des doigts est connue. On va donc garder les informations obtenues en mémoire et lorsqu'une nouvelle information de main arrive du module Leap Motion, on identifie dans un premier temps la position des doigts et le geste effectué. Si ce geste est identique aux gestes précédents enregistrés, on évalue la position de la main et on voit s'il y a eu mouvement de celle-ci entre les différentes prises. On obtient alors les valeurs de déplacement. Si celles-ci se trouvent dans un range de déplacement défini alors, ce mouvement de la main est accepté.

Enfin, dans la configuration de l'action sélectionnée, soit on limite l'action à une opération à la fois (répondre au téléphone par exemple), soit on permet une action continue. tant que la position de la main ne change pas au moment où le geste a été validé. Lorsque que le geste n'est plus reconnu, l'action s'arrête. (ex : ouverture des vitres)

## 11.4 Différences entre main gauche et main droite

Si on définit un range de valeurs pour chaque doigt, on peut facilement définir une reconnaissance efficace pour chaque geste. Cette technique peut se montrer assez efficace. En remarquant le fait que si les actions choisies pour la main droite et la main gauche sont identiques, on pourrait penser proposer les mêmes range de valeurs pour détecter un geste de la main droite et de la main gauche, ce qui permettrait de faire gagner du temps pour chaque geste. En réalité, il n'en est rien.

Le module Leap Motion fait la différence entre main gauche et main droite et donc, peut décoder les informations des différents doigts, ce qui fait que le premier doigt perçu dans les informations renvoyées est le pouce et le dernier l'auriculaire. Mais les valeurs données diffèrent. L'exemple le plus frappant est le pouce : si on ouvre tous les doigts des deux mains, les pouces auront une direction opposée. D'autres modifications de valeurs se présentent suite à cette différence. Au final ces différences font que si on reconnaît un geste

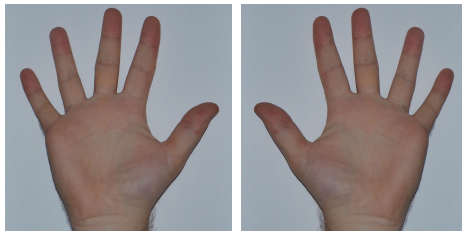


FIGURE 11.5 – différence main gauche et main droite

de la main gauche, il faut adapter les ranges de valeurs autorisés pour que le geste soit reconnu correctement. Une possibilité pour éviter ce problème serait, lors du calcul des positions des doigts de la main, d'effectuer une opération de miroir d'une des deux mains, de façon à ce que si celle-ci fait le même geste, elle puisse être reconnue de la même façon. L'idée n'a pas été retenue car cela complexifie l'étude des valeurs à employer pour la reconnaissance des geste.

## 11.5 Stockage des valeurs

Un certain nombre d'informations de mouvement doivent être stockées. Dans le cas présent, une base de donnée MYSQL se chargera du travail. Ce système offre l'avantage de permettre de stocker, de modifier et de récupérer les informations de gestuelle aisément. Sur un système de type PC, ce choix semble fort logique.

Si une telle technologie devait être intégrée dans un véhicule, la logique de stockage des informations serait probablement différente. Il est clair que les technologies embarquées dans un véhicule ne respectent pas le même cahier des charges. Ce choix d'une base de donnée MYSQL dans ce cas ne sera pas forcément une bonne idée dans de telles conditions. Il sera probablement plus logique de s'adapter au véhicule et par exemple, exploiter le système de stockage de données de l'infotainment.

## 11.6 Implémentation du logiciel

Le système Leap Motion est fourni avec des bibliothèques et des logiciels de test. En JAVA, ces bibliothèques sont assez aisées à employer. Des exemples d'utilisation sont également proposés pour bien comprendre le fonctionnement. Le problème est que les exemples proposés fournissent les données de façon brute (Fig. 11.6), ce qui ne permet pas de comprendre aisément ce que voit le module. Dès lors, lorsqu'un geste est effectué, s'il n'est pas reconnu, on se pose la question de l'origine du problème. Est-ce dû à un mauvais range de valeurs, le module a-t-il mal perçu un élément ou bien y a-t-il une erreur d'interprétation? Plus largement, on en arrive rapidement à se demander ce que voit le capteur. Pour savoir ce qu'il voit, exploiter les données brutes est trop complexe, il faut faciliter la reconnaissance. Pour ce faire,

```
Frame id: 11052, timestamp: 1604075421348081, hands: 1, fingers: 5, tools: 0, gestures 1
Left hand, id: 1, palm position: (1.16753, 130.841, 101.862)
pitch: -49.53873367269884 degrees, roll: 16.11055331796396 degrees, yaw: 67.99268903640103 degrees
Arm direction: (0.87317, 0.484206, -0.0558506), wrist position: (-47.5495, 147.989, 113.629), elbow position: (-275.823, 21.4022, 128.23)
TYPE THUMB, id: 10, length: 50.462393mm, width: 19.687336mm
TYPE METACARPAL bone, start: (-34.8441, 166.255, 133.306), end: (-34.8441, 166.255, 133.306), direction: (0, 0, 0)
TYPE PROXIMAL bone, start: (-34.8441, 166.255, 133.306), end: (12.86, 172.942, 129.341), direction: (-0.986979, -0.138359, 0.082028)
TYPE INTERMEDIATE bone, start: (12.86, 172.942, 129.341), end: (45.1499, 166.698, 126.468), direction: (-0.978082, 0.189151, 0.0870435)
TYPE DISTAL bone, start: (45.1499, 166.698, 126.468), end: (67.7053, 165.71, 124.52), direction: (-0.995345, 0.0435997, 0.0859534)
TYPE INDEX, id: 11, length: 56.94117mm, width: 18.728928mm
TYPE METACARPAL bone, start: (-32.5807, 164.709, 110.518), end: (33.8325, 141.207, 99.9669), direction: (-0.932314, 0.329924, 0.148122)
TYPE PROXIMAL bone, start: (33.8325, 141.207, 99.9669), end: (66.6358, 121.406, 116.164), direction: (-0.788561, 0.475987, -0.389368)
TYPE INTERMEDIATE bone, start: (66.6358, 121.406, 116.164), end: (80.587, 109.453, 130.662), direction: (-0.59612, 0.51076, -0.619488)
TYPE DISTAL bone, start: (80.587, 109.453, 130.662), end: (87.8871, 101.041, 142.895), direction: (-0.441273, 0.508459, -0.739424)
TYPE MIDDLE, id: 12, length: 64.87989mm, width: 18.394297mm
TYPE METACARPAL bone, start: (-36.3069, 155.24, 104.306), end: (22.7378, 125.304, 90.8491), direction: (-0.874039, 0.443144, 0.1992)
TYPE PROXIMAL bone, start: (22.7378, 125.304, 90.8491), end: (56.4684, 100.833, 111.863), direction: (-0.722738, 0.524328, -0.450255)
TYPE INTERMEDIATE bone, start: (56.4684, 100.833, 111.863), end: (70.6253, 87.8951, 131.619), direction: (-0.514161, 0.469878, -0.717532)
TYPE DISTAL bone, start: (70.6253, 87.8951, 131.619), end: (77.8353, 80.4802, 146.949), direction: (-0.35228, 0.40751, -0.842517)
TYPE RING, id: 13, length: 62.383747mm, width: 17.503330mm
TYPE METACARPAL bone, start: (-41.3576, 145.041, 100.844), end: (6.61403, 111.115, 85.7981), direction: (-0.798933, 0.559364, 0.248067)
TYPE PROXIMAL bone, start: (6.61403, 111.115, 85.7981), end: (37.3079, 86.4963, 103.781), direction: (-0.709494, 0.569059, -0.415681)
TYPE INTERMEDIATE bone, start: (37.3079, 86.4963, 103.781), end: (51.8502, 73.5016, 122.196), direction: (-0.54216, 0.484464, -0.686555)
TYPE DISTAL bone, start: (51.8502, 73.5016, 122.196), end: (59.2122, 66.1841, 137.013), direction: (-0.406944, 0.404483, -0.819817)
TYPE PINKY, id: 14, length: 48.907703mm, width: 15.547833mm
TYPE METACARPAL bone, start: (-48.195, 133.513, 104.046), end: (-9.09489, 97.5844, 85.8095), direction: (-0.696413, 0.639925, 0.324815)
TYPE PROXIMAL bone, start: (-9.09489, 97.5844, 85.8095), end: (10.4033, 71.7618, 96.9983), direction: (-0.569507, 0.754229, -0.326802)
TYPE INTERMEDIATE bone, start: (10.4033, 71.7618, 96.9983), end: (20.412, 60.2464, 108.217), direction: (-0.528496, 0.608059, -0.592416)
TYPE DISTAL bone, start: (20.412, 60.2464, 108.217), end: (28.3623, 52.1578, 120.461), direction: (-0.476357, 0.484645, -0.733623)
Key Tap id: 1, STATE_STOP, position: (49.7121, 151.903, 145.399), direction: (-0.0129094, -0.11686, 0.993065)
```

FIGURE 11.6 – Informations d'une main dans le programme d'exemple de Leap Motion

une fenêtre de debugging a été intégrée au programme dans le but de voir ce qui se passe. Cette fenêtre doit permettre de présenter les informations pour les rendre aisément décodables des informations brutes exploitées par le programme mais également de comprendre le décodage qui a été effectué et les informations qui en résultent.

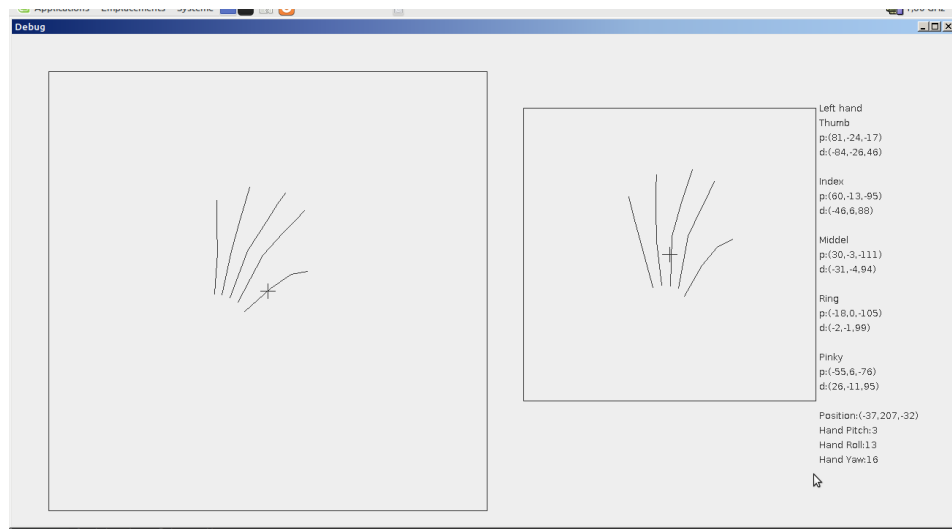


FIGURE 11.7 – Informations d’une main dans le programme d’exemple de Leap Motion

Cela doit au final permettre de comprendre ce que voit le programme et d’identifier aisément les problèmes qui se présentent.

Dans le cadre de gauche (Fig. 11.7), on peut voir un dessin de la main tel qu’il est vu par le module. On peut remarquer qu’il n’est pas centré ni aligné avec le centre du cadre car la main dont la position a été enregistrée n’est pas centrée face au module lors de la capture. Chaque ligne dessinée correspond à un os de la main tel que dessiné sur l’image (Fig. 11.3).

A noter que sur cette image, l’axe de profondeur est volontairement ignoré, seul les axes X et Z sont pris en compte. La raison en est que pour afficher les trois axes, il aurait fallu reporter les lignes dans une image en trois dimensions. Cependant, l’objectif de cette fenêtre est clairement de faciliter la compréhension de ce que voit le programme. Il est cependant évident que si l’axe de profondeur est ignoré sur le dessin, il est bien sûr pris en compte au même titre que les autres axes dans les calculs d’identification de la gestuelle.

Ainsi que cela a été expliqué plus tôt, le programme va chercher à re-centrer la main et à extraire les valeurs de position et de rotation. Le centre de la main reçoit donc la position 0,0,0 sans rotation. Tous les autres points de la main sont recalculés en conséquence. Cela donne l’image du cadre de droite. Les informations de cette image sont indiquées en chiffres à droite du cadre. Ce sont ces informations qui sont prises en compte et évaluées par rapport aux ranges qui ont été définis pour chaque geste afin de permettre d’identifier la gestuelle.



Cette fenêtre est très utile pour comprendre si le code de décodage présente un bug, s'il y a une erreur au niveau des ranges de valeurs pour les gestes à reconnaître, ou tout autre problème. Au niveau du module Leap Motion, cela montre aussi son niveau d'efficacité. Il arrive souvent qu'il interprète mal la main qu'il voit, qu'il reconnaisse et représente une main gauche alors que c'est une main droite qu'il voit ou qu'il se trompe dans l'interprétation des doigts de la main en les positionnant mal. De plus, même lorsqu'il décode correctement l'image reçue, les informations montrent souvent une imprécision qui peut empêcher de reconnaître certains gestes. Ce problème sera discuté un peu plus loin dans ce document.

## 11.7 Feedback sonore au conducteur

La sécurité est un élément central dans le développement. L'objectif de ce programme est de permettre au conducteur d'effectuer une action sans avoir à quitter la route des yeux car il a été prouvé que cela avait un impact critique sur l'efficacité de la conduite[55, 17, 35]. Mais les recherches effectuées sur ce système montrent que la charge mentale augmente significativement dès lors qu'aucun feedback de l'action demandée n'est transmis au conducteur[55].

Lors du développement du programme, le problème a été vite ressenti et décision a été prise d'ajouter cette fonctionnalité. Partant du principe que la fenêtre est une fenêtre de debugging et qu'elle accapare l'attention visuelle, il est évident qu'au-delà du développement, elle ne sera plus accessible.

Au sein d'une voiture en condition réelle ou même déjà en phase d'évaluation du système, elle n'a pas sa place. Il faut proposer un retour qui permette au conducteur d'être sûr que l'action a été reçue par le programme et que l'action reçue est bien celle demandée par le conducteur. Pour cela, l'une des solutions proposées dans les documents étudiés est l'utilisation d'un retour sonore [55, 44, 46, 35]. Très efficace, il permet au conducteur de recevoir un feedback sans avoir à quitter la route des yeux.

On va donc proposer de jouer un son. S'il est joué par l'ordinateur dans l'évaluation, en condition réelle, il le serait par le système d'infotainment du véhicule. Pour permettre au conducteur de s'assurer que le geste a été correctement interprété, le nom du son à jouer est enregistré dans les données du geste dans la base de données du système. Cette idée est déjà proposée dans le document [46]. Cela permet de définir quel son jouer pour quelle action et donc d'avoir des sons différents. Ces sons peuvent être de différents types, simple sonnerie, voix, etc... Ce son peut être joué une seule fois ou de



façon répétée si on effectue une action continue.

## 11.8 Evaluation et test du système

Une fois le programme écrit et fonctionnel, les différents gestes qui ont été choisis et décrits précédemment ont été étudiés pour en extraire des ranges de valeurs pour leur reconnaissance et ces informations ont été introduites dans la base de données, comme illustré sur l'image (Fig. 11.8), baisser la vitre et baisser le volume.

A partir de là, on peut remarquer que le module indique une action différente avec deux gestes parfaitement identiques mais effectués une fois avec la main gauche et une fois avec la main droite. La fenêtre de debugging indique clairement qu'il a identifié le geste mais aussi la main qui l'a effectué et a agi en conséquence (Fig. 11.8). Notons que sur la gestuelle employée, on fonctionne sur la logique d'une voiture avec le volant à gauche. La main gauche commande les vitres et la droite commande l'infotainment (Fig. 11.8). Dans le cas présent, un seul module Leap Motion reçoit les commandes des deux mains mais dans un véhicule, on exploiterait deux modules différents positionnés de part et d'autre du conducteur.

Pour effectuer un test correct et efficace du système il va falloir deux choses :

1. un panel de testeurs suffisamment large (50 personnes par exemple) qui devront effectuer différentes actions avec la main droite et la main gauche et évaluer la réponse du programme en fonction de l'action demandée.
2. définir un point de mesure pour évaluer l'efficacité du système.

Cela pose un problème : la situation sanitaire actuelle ne permet pas de réunir de quelque façon que ce soit les personnes pour leur faire effectuer les tests souhaités.

Comme point de mesure, on aurait pu évaluer le programme sur le taux de reconnaissance des gestes demandés. Comme vu précédemment, les tests se font généralement sur deux points :

1. objectif : évaluer le taux de réussite d'un éventail d'opérations demandées aux testeurs
2. subjectif : évaluer le ressenti des utilisateurs via un questionnaire du type de ceux présentés précédemment.

A ce stade, en prenant en compte le fait qu'aucune phase de test correcte n'a été effectuée, le programme semble montrer qu'il est capable de reconnaître et de gérer des gestuelles identiques pour les deux mains et d'agir en fonction de la main qui a effectué le geste. C'est l'élément central de ce document. Les

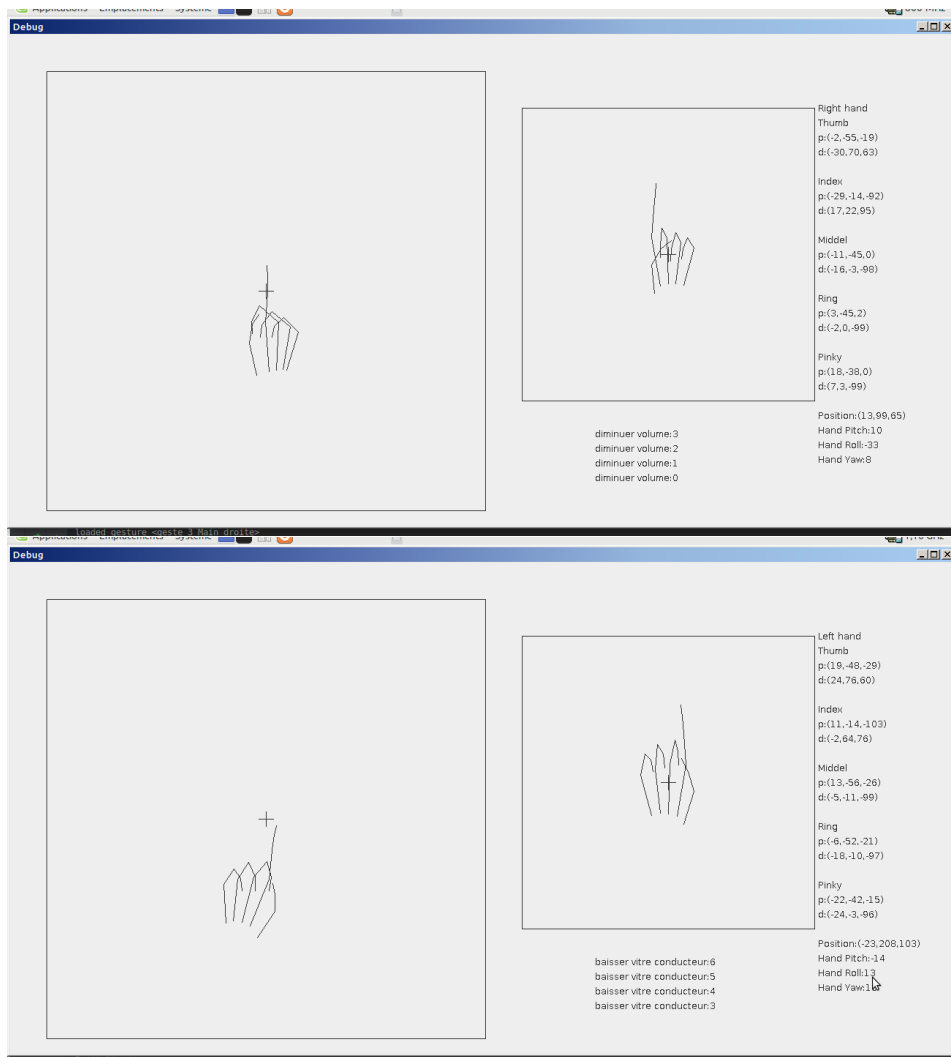


FIGURE 11.8 – Exemple de gestuelle encodée et reconnue : baisser la vitre conducteur avec la main gauche et diminuer le volume sonore avec la main droite

différentes actions demandées au programme ont prouvé qu'il est possible d'identifier la main et de proposer des actions différentes.

Enfin si les tests avec le panel d'utilisateurs avait été concluant, un autre test avec une phase en simulateur aurait pu être envisagé pour se diriger ensuite vers un test en conditions réelles.

## 11.9 Efficacité et précision du système.

Si à ce stade, le testing est difficile, on peut cependant faire des suppositions sur base des documents étudiés, sur les choix qui ont été décidés et se livrer à une première évaluation.

### 11.9.1 Module "Leap Motion"

S'il a montré qu'il était capable de reconnaître les gestes effectués dans la plupart des cas, le système Leap Motion n'est pas un système sans défaut. Fréquemment, il se montre incapable de reconnaître une main de façon correcte. Un problème assez fréquent, est le fait qu'il interprète mal les informations qui lui sont données. Sur l'image (Fig. 11.9), l'écran représente une main gauche alors que devant le capteur, c'est une main droite qui est représentée. Si on observe bien l'image, sur les lignes de la main qui est représentée à l'écran, le pouce se trouve clairement à droite, alors qu'il devrait se positionner à gauche. Une technique possible pour éviter un tel problème en condition réelle serait de limiter la reconnaissance d'une seule main par capteur. Chaque capteur ne pourrait provoquer que les actions possibles par la gestuelle de la main qu'il est sensé gérer. Même lorsque la main est correctement reconnue, il arrive que les doigts soient mal positionnés. Le problème d'une telle situation est qu'on peut en arriver à percevoir un geste qui n'est pas le bon et déclencher une action qui n'est pas souhaitée. Sur l'image, le geste présenté devant le capteur montre l'index levé seulement, mais sur l'image de la fenêtre, on voit que c'est le majeur est levé lui aussi (Fig. 11.10). Un dernier problème est la précision du système sur l'image (Fig. 11.11). Le geste présenté devant le module est un toucher pouce-index. Sur la fenêtre de debug, on voit qu'il existe une distance assez importante entre le pouce et l'index. On peut remarque également que la distance entre l'index et le majeur est plus faible que dans la réalité. Cela signifie que pour que le système reconnaisse le toucher pouce-index, il faut donner un range qui permette une distance suffisamment grande pour que cette différence n'empêche pas la détection du geste effectué. Mais si la distance tolérée est trop importante, sur cette image, le geste toucher pouce-majeur sera également détecté, ce qui n'est pas le geste souhaité. Les gestes de toucher entre le pouce et un autre doigt est fortement impacté par cette imprécision du module de reconnaissance. Les autres gestes (1,2,3,4 par exemple) le sont beaucoup moins, car devant le capteur, la position de chaque doigt diverge plus d'un geste à l'autre.

Au final, il est évident que pour développer et tester un système de reconnaissance de gestuelle des deux mains, ce système peut se montrer assez efficace pour peu qu'on soit un peu patient et qu'on ne lui présente pas de gestuelle trop complexe. Néanmoins les erreurs qu'il peut faire et son

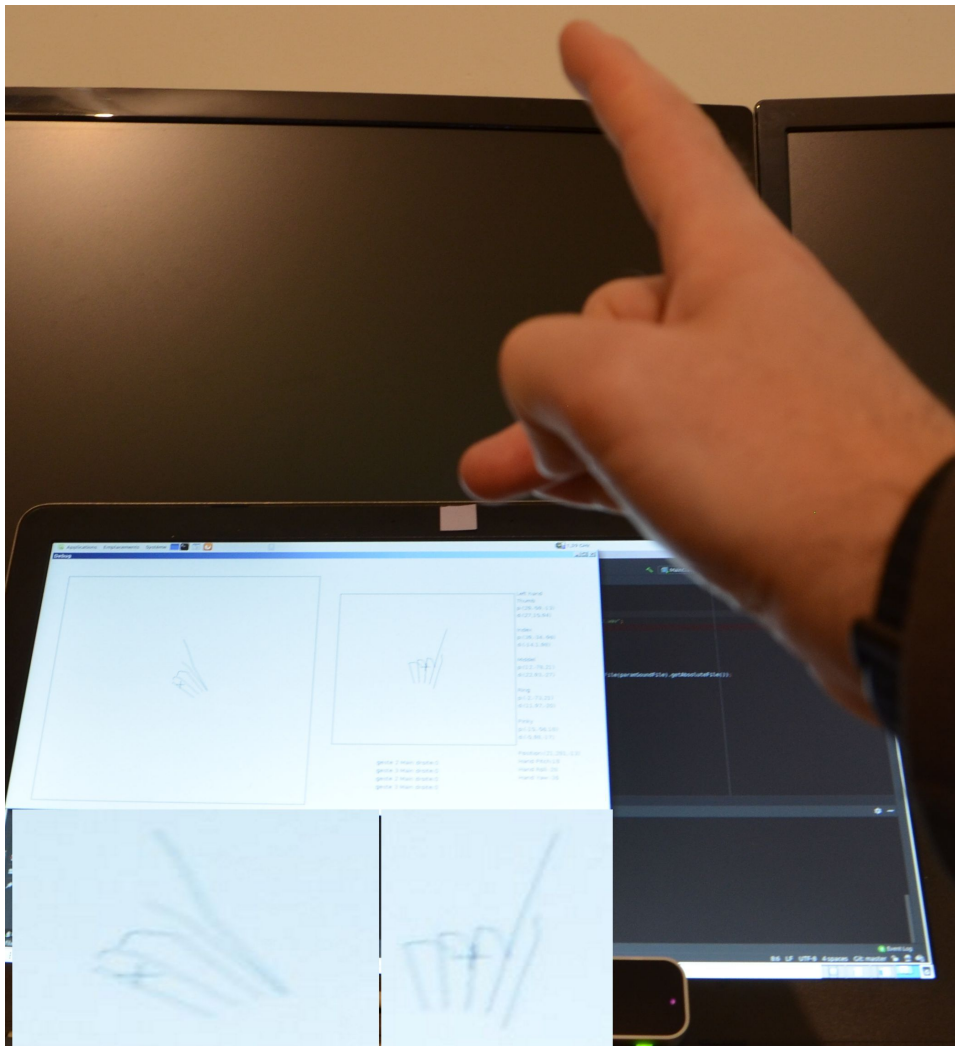


FIGURE 11.9 – Main mal reconnue, main gauche représentée à l'écran mais main droite devant le capteur

manque de précision rendent son utilisation en conditions réelles au sein d'un véhicule peu envisageable. De plus, même pour une phase de test, il faudra prendre en compte une telle imprécision. L'utilisation d'un module différent pourrait donner des résultats différents avec plus ou moins d'erreurs.

Le programme développé n'est pas non plus exempt de défauts. Certains gestes qui ont été proposés plus haut ne sont pas possibles à cause de la façon dont les données sont enregistrées. De plus par rapport au capteur, les ranges de valeurs tolérées pour les capteurs ont été évaluées de façon arbi-

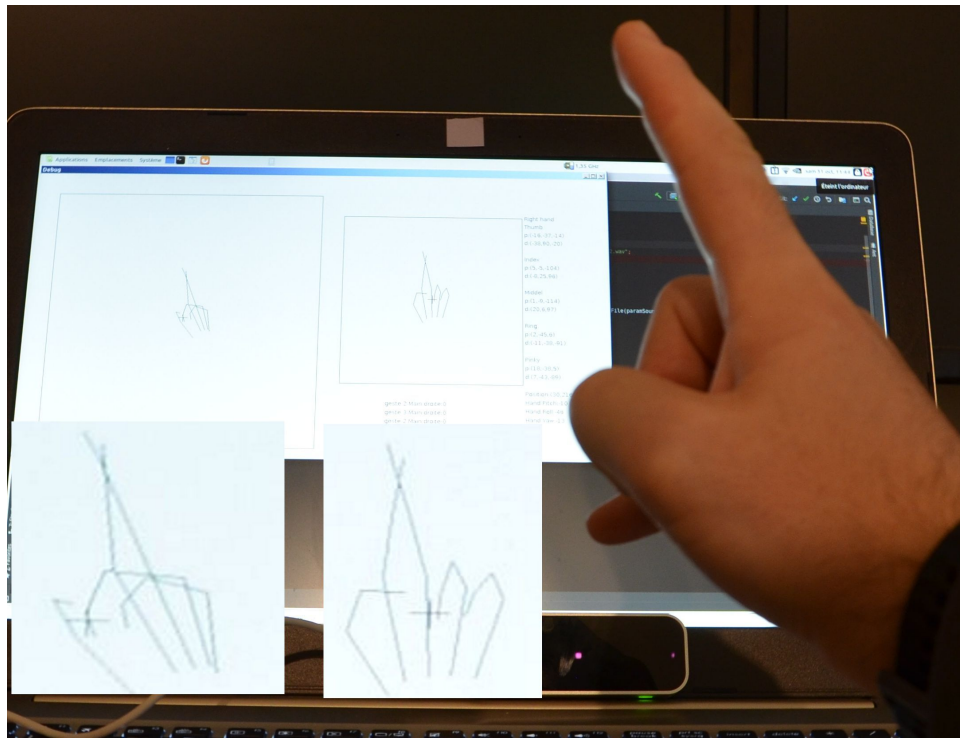


FIGURE 11.10 – Main mal reconnue, le majeur est baissé devant le capteur mais il est indiqué comme étant levé

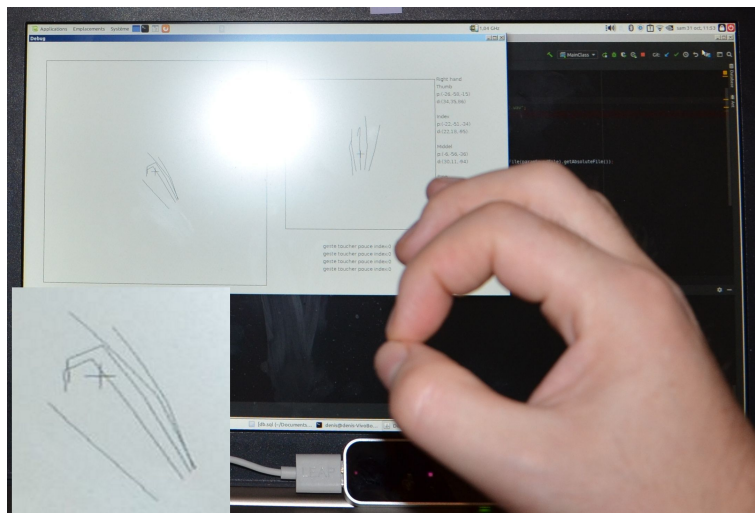


FIGURE 11.11 – Manque de précision de la reconnaissance pouce-index

traire. Des recherches plus poussées auraient dû être effectuées en prenant une image des mains d'un panel de personnes effectuant différents gestes

pour évaluer les meilleures valeurs afin de maximiser l'efficacité du système.

## 11.10 Conséquences des problèmes de tolérances

Ces problèmes de précision du système posent la question de la précision et des erreurs du système. Doit-on permettre des ranges de valeurs plus larges au risque d'avoir plus de détections de faux positifs avec pour conséquence des actions non désirées? Au contraire, faut-il les limiter au maximum et accepter le fait que certains gestes pourtant correctement exécutés ne seront pas perçus, ce qui augmenterait la charge mentale du conducteur et son niveau de stress. Dans les deux cas, on en arriverait rapidement à une mauvaise perception du système.

## 11.11 Sécurité

On ne peut développer un tel module sans prendre en compte l'aspect de la sécurité. Les problèmes que peut présenter le module montrent qu'il est important de développer un système efficace capable de reconnaître les gestes sans faire d'erreur. Une gestuelle mal perçue ou pas perçue du tout provoquerait une augmentation du stress et une perte d'attention pour l'activité première, le conducteur cherchant à identifier la raison du dysfonctionnement du système de perception ce qui conduirait à diminuer son attention et le pousserait à quitter la route des yeux. Au final, alors que le système se veut plus sécurisant, il perd son intérêt et le résultat est pire que mieux.

Un autre aspect de la sécurité est le problème du piratage. Si la zone de perception d'un module est trop large, il pourrait être exploité pour effectuer des actions contre l'avis du conducteur. Si le module côté porte peut percevoir une zone trop large, positionner sa main sur la vitre dans l'axe du capteur pourrait permettre à une personne mal intentionnée d'ouvrir les vitres du véhicule, mettre le conducteur en danger ou faciliter un car-jacking.

Le capteur central n'est pas en reste. Un geste effectué sur le pare-brise pourrait permettre là aussi une action dangereuse si la configuration du système permettait la gestion d'éléments critiques du véhicule.

Il est donc vital de définir une zone de reconnaissance de la gestuelle pour chaque capteur : elle doit être précise et bien positionnée dans le véhicule et doit gérer le capteur de façon à ce que en dehors de cette zone, aucun mouvement ne soit reconnu.

## Chapitre 12

# Conclusion

### 12.1 Etat de l'art

Les différents documents qui ont été récoltés montrent plusieurs choses : l'objectif principal des interfaces de communication avec le conducteur dans le véhicule ont pour but d'améliorer la sécurité. Les recherches ont démontré que les conducteurs cherchent de plus en plus à faire d'autres choses que conduire leur véhicule[18]. La conduite du véhicule doit rester la tâche première, toute autre tâche doit passer au second plan et doit donc limiter les besoins physiques et la charge mentale au minimum. Si certains sens du conducteur peuvent être plus facilement exploités pour cette seconde tâche, tels que le toucher, l'audition ou la gestuelle, d'autres doivent être exploités intelligemment voire être évités autant que possible. Ainsi en est-il de la vue du conducteur car l'impact est très important sur la qualité de la conduite.

Ce fait apparaît fortement dans les document étudiés. La sécurité est souvent le but recherché. Dans ces conditions, les interfaces sont étudiées spécifiquement dans ce but. On en arrive pour cela à faire des comparatifs des différentes interfaces dans différentes situations. L'objectif est de tester quelles sont les meilleures interfaces dans les situations que l'on souhaite étudier.

La plupart des recherches ne peuvent s'effectuer en conditions réelles. Les conditions d'utilisation des interfaces sont plus aisées en laboratoire qu'en conditions réelles, celles-ci se montrent souvent beaucoup plus efficaces mais peuvent ensuite se montrer décevantes en conditions réelles. Cependant, ces tests n'en sont pas moins très utiles. Ils permettent en effet de classer les interfaces et la façon de les utiliser qui pourraient se montrer les plus efficaces en conditions réelles. Si une interface montre déjà une efficacité limitée voire se révèle dangereuse, il est très probable qu'elle ne fera pas mieux en conditions réelles.

Un autre élément est qu'on en arrive aujourd'hui à exploiter ensemble ces interfaces. Une première façon de procéder est de combiner ces interfaces dans le but de savoir quelle interface "conducteur vers véhicule" se marie le mieux avec l'interface de réponse. Une fois encore, c'est la sécurité qui est l'élément central. Par exemple, si une interface cherche à limiter l'impact visuel, on va chercher une interface de réponse qui limite l'impact tout autant sous peine de perdre le bénéfice sur la sécurité.

Une autre façon de procéder est d'étudier la possibilité de les exploiter ensemble. Peut-on combiner différentes interfaces et quel est l'impact sur le conducteur ? La sécurité est-elle améliorée par le fait que le conducteur peut recevoir une information par plusieurs canaux et ignorer les autres, ou cette avalanche d'information provoque-t-elle une surcharge mentale importante pour le conducteur ? Les documents étudiés montrent qu'on se trouve un peu entre les deux.

Un dernier élément à noter est que si la sécurité est l'élément central des interfaces de communication avec le conducteur dans le véhicule, l'arrivée des véhicules totalement autonomes sans conducteur change la donne. Dans cette situation spécifique, les interfaces peuvent exploiter toute l'attention sans risque. A noter que certaines interfaces peuvent également le faire dans des véhicules fortement automatisés pour peu qu'on puisse les rendre inutilisables en conduite manuelle.

## 12.2 Recherche

Le programme utilisé avec le module a montré qu'il était possible de définir une gestuelle différente avec les mains et d'agir différemment en fonction de la main qui est présentée devant le capteur. C'est l'objectif à atteindre. Cependant, si le programme a montré que c'était possible, sa conception a mis en évidence ses défauts et soulevé d'autres questions résultant soit de l'utilisation des gestuelles multiples, soit de l'utilisation de cette technologie, ou encore sur les actions possibles que l'on peut en attendre.

Si les gestes proposés sont peu nombreux, l'espace requis par le système se montre peu intéressant face à l'utilisation de boutons mécaniques utilisables par retour tactile sans quitter la route des yeux, pour peu que le conducteur connaisse l'habitacle de son véhicule. Proposer plus de gestes demande au conducteur un apprentissage. Mais ici, l'utilisation des deux mains permet de doubler le nombre de commandes possibles avec une gestuelle limitée.

Les résultats obtenus par le programme et les informations trouvées dans



les documents qui ont étudié les capteurs semblent indiquer qu'on pourrait effectivement utiliser un tel système dans un véhicule, mais il est évident que la technologie et les algorithmes de reconnaissance utilisés devraient montrer une efficacité très élevée pour pouvoir présenter un intérêt réel.

Un dernier problème est celui des actions elles-mêmes. On propose ici, pour la main côté portière, de gérer les commandes des vitres. Mais cette idée est discutable car cela sous-entend qu'un tel système montre des avantages par rapport à un pack classique de boutons comme on en trouve aujourd'hui dans les voitures sur la porte conducteur. Avec ce système classique, le conducteur peut très bien ouvrir et fermer les fenêtres sans avoir à quitter la route des yeux. Dès lors, cela rend discutable l'intérêt de ce système à ce niveau. Cependant on pourrait proposer d'autres gestes, le système de ce côté permettrait de gérer plus d'éléments ce qui rendrait son utilisation intéressante.

Une dernière possibilité serait de permettre au conducteur de programmer ses actions, mais une telle opération prendrait du temps, et rendrait le véhicule moins pratique aux autres personnes qui le conduiraient sans connaître la configuration du système de reconnaissance gestuelle et éprouveraient des difficultés à le manipuler. De plus, le conducteur pourrait programmer des opérations peu gérables en conduite sans s'en rendre compte et le système pourrait se retrouver à effectuer des actions incorrectes voire dangereuses.

## Chapitre 13

# Recherches futures

Bien qu'ayant permis d'avancer quant à la possibilité d'utiliser les interfaces de reconnaissance de gestuelle des deux mains, ce document laisse de nombreux points en suspens sur lesquels il est encore possible de faire des recherches et des améliorations.

### 13.1 Interface

Le choix de l'interface, la façon de l'utiliser, évaluer et améliorer sa précision pourrait faire l'objet de plus de recherches. Le choix d'un autre module ou d'une autre technologie pourrait aussi être étudié. Dans certains documents, d'autres techniques telles que l'utilisation d'un radar ou d'une caméra sont proposés. On peut s'interroger sur les possibilités et des problèmes qu'on rencontrerait en utilisant ces techniques de reconnaissance.

De plus pour faire les tests, on a choisi d'utiliser un module Leap Motion. Mais il existe d'autres systèmes et d'autres technologies. On pourrait faire une review des technologies existantes, et des interfaces qui exploitent ces technologies. On pourrait évaluer quelle interface et quelle technologie se montrerait plus efficace dans un environnement de voiture, en prenant en compte le fait que deux interfaces devraient cohabiter. On pourrait également se poser la question de savoir s'il serait possible avec un autre système que le Leap Motion de limiter à une seule interface la reconnaissance de gestuelle des deux mains, avec une réflexion sur tous les problèmes que cela pose.

### 13.2 Gestuelle

En ce qui concerne la gestuelle, il est possible d'effectuer encore effectuer encore d'autres recherches. La logique utilisée par l'application de reconnaissance est assez simple, mais pour reconnaître un geste, il faut que la position

des doigts se trouve dans un range de valeurs. Or ce range de valeurs a été choisi de façon arbitraire. De plus, aucun test n'a été effectué sur un pool d'utilisateurs. Une suite possible pour ce travail serait donc de regrouper un panel de testeur pour déterminer les valeurs à utiliser pour la reconnaissance qui rendrait le programme le plus efficace possible.

On pourrait également se poser la question l'efficacité de la technique utilisée. Quels avantages pourraient apporter l'utilisation d'une autre technique? En règle générale, on peut se poser la question de l'agoritmique utilisée dans ce cas de figure précis.

Des améliorations sont possibles également dans la gestuelle choisie. Même si les choix de gestes et d'actions ont été faits sur base des informations acquises dans l'étude des autres documents, d'autres solutions sont possibles et rien ne permet de dire que les choix qui ont été fait sont les bons. On pourrait par exemple exploiter le panel de testeurs afin leur proposer de tenter de répondre à plusieurs questions : quel est leur ressenti vis-a-vis d'un système de gestuelle à deux mains, quelles gestuelles préféreraient-ils exécuter? Cette gestuelle choisie serait-elle la même pour les deux mains? Quelle actions souhaitent-ils pouvoir effectuer avec cette gestuelle?

De plus, il a été décidé de proposer des gestes assez proches mais différents selon la main qui a effectué le geste. On peut se demander si ce choix est judicieux. On peut aussi poser la question de la différence de gestuelle des deux mains. Faut-il un gestuelle fortement différente ou doit-on préférer une gestuelle la plus proche possible pour les deux mains?

En ce qui concerne les actions proposées, nous avons décidé de permettre des actions différentes pour les deux mains. On peut là aussi se demander si ce choix est judicieux. Il faudrait évaluer l'efficacité des choix de gestuelles allant d'actions totalement différentes en fonction de la main qui a effectué le geste à des actions totalement identiques et en évaluer l'impact.

Une dernière question qu'on peut se poser est de savoir si les actions choisies pour chaque main sont judicieuses. Pour commander les fenêtres et l'infotainment, il a été décidé d'utiliser la main qui se trouve la plus proche de ce système. Il existe de nombreuses autres possibilités qu'on pourrait gérer avec la gestuelle. Il faudrait faire une review des possibilités d'actions et évaluer le niveau d'intérêt de chacune de ces actions effectuées avec une main plutôt que l'autre.

Pour répondre à de telles questions, en plus de demander au panel de testeurs d'effectuer des gestes pour définir les valeurs pour la reconnaissance de gestes, on pourrait demander de répondre à des questions ou d'effectuer des

actions avec un simulateur dans le but d'évaluer et améliorer ce système. Après avoir effectué différentes actions, les testeurs pourraient répondre à des questionnaires objectifs et subjectifs afin d'évaluer et améliorer le concept d'utilisation des deux mains grâce à un système de gestuelle dans un véhicule.

# Glossaire

## A

**ASR** (Automatic Speech Recognition) - Il permet de reconnaître la voix du conducteur, mais cela nécessite un système de réduction du bruit. 31

## C

**CNN** (Convolutional Neural Network) - Technique de réseau neuronal pouvant être utilisé pour la reconnaissance de gestuelle par une caméra. 26

## D

**DDNN** (Distributed Deep Neural Network) - Réseaux neuronaux multicouches pouvant être utilisés pour la reconnaissance de gestuelle par une caméra. 26

## H

**HUD** (Head Up Display) - Afficheur transparent ou sur le pare-brise permettant d'afficher des informations tout en gardant les yeux dirigés vers la route. 38–40, 43, 47

## L

**Leap Motion** Module de reconnaissance de gestuelle des mains. Il permet de différencier main gauche et main droite et de reconnaître les différents éléments qui composent la main pour identifier sa position avec une certaine précision. 24, 59, 61, 63–65, 67, 70, 77

**LSTM** (Long Short Term Memory) - Technique RNN pouvant être utilisée pour la reconnaissance vocale ou la reconnaissance de gestuelle par une caméra. 26

**LTT** (Look To Talk) - Système basé sur le regard permettant d'activer l'écoute d'une commande vocale. 45

## P

**PTT** (Push To Talk) - Bouton servant à activer la reconnaissance vocale dans le véhicule. Il permet d'empêcher une activation intempestive. 30, 45

## **S**

**Sonification** Transmission d'un signal audio non vocal pour transmettre une information. 36

## **T**

**TTS** (Text To Speech) - Système de synthèse vocale qui permet de convertir un texte en signal vocal. 36

# Bibliographie

- [1] B. I. Ahmad, P. M. Langdon, S. J. Godsill, R. Donkor, R. Wilde, and L. Skrypchuk. You do not have to touch to select : A study on predictive in-car touchscreen with mid-air selection. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Automotive'UI 16, page 113–120, New York, NY, USA, 2016. Association for Computing Machinery.
- [2] L. Angelini, J. Baumgartner, F. Carrino, S. Carrino, M. Caon, O. A. Khaled, J. Sauer, D. Lalanne, E. Mugellini, and A. Sonderegger. A comparison of three interaction modalities in the car : Gestures, voice and touch. In *Actes de La 28ième Conference Francophone Sur l'Interaction Homme-Machine*, IHM '16, page 188–196, New York, NY, USA, 2016. Association for Computing Machinery.
- [3] P. Biswas, G. Prabhakar, J. Rajesh, K. Pandit, and A. Halder. Improving eye gaze controlled car dashboard using simulated annealing. In *Proceedings of the 31st British Computer Society Human Computer Interaction Conference*, HCI '17, Swindon, GBR, 2017. BCS Learning ; Development Ltd.
- [4] M. Braun, J. Li, F. Weber, B. Pfleging, A. Butz, and F. Alt. What if your car would care? exploring use cases for affective automotive user interfaces. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '20, New York, NY, USA, 2020. Association for Computing Machinery.
- [5] B. Bressollette, S. Denjean, V. Roussarie, M. Aramaki, S. Ystad, and R. Kronland-Martinet. Harnessing audio in auto control : The challenge of sonifying virtual objects for gesture control of cars. *IEEE Consumer Electronics Magazine*, 7(2) :91–100, 2018.
- [6] P. Di Campli San Vito, S. A. Brewster, F. Pollick, and S. White. Thermal in-car interaction for navigation. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI '17, page 510–511, New York, NY, USA, 2017. Association for Computing Machinery.
- [7] D. Dmitrenko, E. Maggioni, and M. Obrist. Towards a framework for validating the matching between notifications and scents in olfactory

- in-car interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, page 1–6, New York, NY, USA, 2019. Association for Computing Machinery.
- [8] E. Gentet, B. David, S. Denjean, G. Richard, and V. Roussarie. Speech intelligibility enhancement by equalization for in-car applications. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6934–6938, 2020.
  - [9] R. Haeuslschmid, B. Pfleging, and F. Alt. A design space to support the development of windshield applications for the car. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 5076–5091, New York, NY, USA, 2016. Association for Computing Machinery.
  - [10] J. Huber, M. Sheik-Nainar, and N. Matic. Towards an interaction language for force-enabled touchpads in cars. In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '16 Adjunct, page 197–202, New York, NY, USA, 2016. Association for Computing Machinery.
  - [11] U. Inc. <https://developer-archive.leapmotion.com/documentation/csharp/devguide/Leap-Overview.html>.
  - [12] K. Keller, K. V. Carl, H. Jöntgen, B. M. Abdel-Karim, M. Mühlhäuser, and O. Hinz. "k.i.t.t., where are you?" : Why smart assistance systems in cars enrich people's lives. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '19 Adjunct, page 1120–1132, New York, NY, USA, 2019. Association for Computing Machinery.
  - [13] T. Kujala and H. Grahn. Visual distraction effects of in-car text entry methods : Comparing keyboard, handwriting and voice recognition. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '17, page 1–10, New York, NY, USA, 2017. Association for Computing Machinery.
  - [14] M. Lakier, L. E. Nacke, T. Igarashi, and D. Vogel. Cross-car, multi-player games for semi-autonomous driving. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '19, page 467–480, New York, NY, USA, 2019. Association for Computing Machinery.
  - [15] S. Langlois, T. N. That, and P. Mermillod. Virtual head-up displays for augmented reality in cars : A user testing to validate the congruence. In *Proceedings of the European Conference on Cognitive Ergonomics*,



- ECCE '16, New York, NY, USA, 2016. Association for Computing Machinery.
- [16] D. R. Large, L. Clark, G. Burnett, K. Harrington, J. Luton, P. Thomas, and P. Bennett. "it's small talk, jim, but not as we know it." : Engendering trust through human-agent conversation in an autonomous, self-driving car. In *Proceedings of the 1st International Conference on Conversational User Interfaces*, CUI '19, New York, NY, USA, 2019. Association for Computing Machinery.
  - [17] J. Mahetalia, R. Mehta, S. Varudandi, and K. Samdani. A comparative study on gesture recognition systems for car infotainment systems. In *2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA)*, pages 1–7, 2018.
  - [18] N. Martelaro, J. Teevan, and S. T. Iqbal. An exploration of speech-based productivity support in the car. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–12, New York, NY, USA, 2019. Association for Computing Machinery.
  - [19] M. Martin, F. Geiger, M. Götz, T. Beeh, M. Sosnowski, M. Keppner, I. Aslan, B. Bittner, and E. André. Traeddy : A stress sensitive traffic jam companion for car commuters. In *Proceedings of the Workshop on Human-Habitat for Health (H3) : Human-Habitat Multimodal Interaction for Promoting Health and Well-Being in the Internet of Things Era*, H3 '18, New York, NY, USA, 2018. Association for Computing Machinery.
  - [20] K. Matsumura and D. S. Kirk. On active passengering : Supporting in-car experiences. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(4), Jan. 2018.
  - [21] A. Matviienko, A. Löcken, A. El Ali, W. Heuten, and S. Boll. Navilight : Investigating ambient light displays for turn-by-turn navigation in cars. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '16, page 283–294, New York, NY, USA, 2016. Association for Computing Machinery.
  - [22] S. Mayer, H. V. Le, A. Nesti, N. Henze, H. H. Bühlhoff, and L. L. Chuang. The effect of road bumps on touch interaction in cars. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '18, page 85–93, New York, NY, USA, 2018. Association for Computing Machinery.
  - [23] A. Meschtscherjakov, A. Krischkowsky, K. Neureiter, A. Mirnig, A. Baumgartner, V. Fuchsberger, and M. Tscheligi. Active corners : Collaborative in-car interaction design. In *Proceedings of the 2016 ACM*

- Conference on Designing Interactive Systems*, DIS '16, page 1136–1147, New York, NY, USA, 2016. Association for Computing Machinery.
- [24] T. Miura, K.-i. Yabu, Y. Shimizu, K. Tanaka, M. Furukawa, S. Michiyoshi, T. Yamamoto, K. Ueda, and T. Ifukube. Cognitive workload of in-car auditory-vocal interfaces on visuospatial sketchpad based on a dual task of visual pattern test. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, SAC '19, page 2319–2327, New York, NY, USA, 2019. Association for Computing Machinery.
  - [25] N. Müller, B. Eska, R. Schäffer, S. T. Völkel, M. Braun, G. Wiegand, and F. Alt. Arch'n'smile : A jump'n'run game using facial expression recognition control for entertaining children during car journeys. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*, MUM 2018, page 335–339, New York, NY, USA, 2018. Association for Computing Machinery.
  - [26] K. Nathwani, M. Daniel, G. Richard, B. David, and V. Roussarie. Formant shifting for speech intelligibility improvement in car noise environment. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5375–5379, 2016.
  - [27] R. Neßelrath, M. M. Moniri, and M. Feld. Combining speech, gaze, and micro-gestures for the multimodal control of in-car functions. In *2016 12th International Conference on Intelligent Environments (IE)*, pages 190–193, 2016.
  - [28] A. Ng and S. Brewster. An evaluation of touch and pressure-based scrolling and haptic feedback for in-car touchscreens. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '17, page 11–20, New York, NY, USA, 2017. Association for Computing Machinery.
  - [29] A. Ng and S. A. Brewster. Investigating pressure input and haptic feedback for in-car touchscreens and touch surfaces. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Automotive'UI 16, page 121–128, New York, NY, USA, 2016. Association for Computing Machinery.
  - [30] A. Ng, S. A. Brewster, F. Beruscha, and W. Krautter. An evaluation of input controls for in-car interactions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 2845–2852, New York, NY, USA, 2017. Association for Computing Machinery.
  - [31] S. Oviatt. Ten myths of multimodal interaction. *Commun. ACM*, 42(11) :74–81, Nov. 1999.
  - [32] A. Panda. Denoising algorithms using stacked rnn models for in-car speech recognition system. In *2018 4th International Conference for Convergence in Technology (I2CT)*, pages 1–5, 2018.

- [33] I. Politis, P. Langdon, D. Adebayo, M. Bradley, P. J. Clarkson, L. Skrypchuk, A. Mouzakitis, A. Eriksson, J. W. H. Brown, K. Revell, and N. Stanton. An evaluation of inclusive dialogue-based interfaces for the takeover of control in autonomous cars. In *23rd International Conference on Intelligent User Interfaces, IUI '18*, page 601–606, New York, NY, USA, 2018. Association for Computing Machinery.
- [34] S. D. Regani, Q. Xu, B. Wang, M. Wu, and K. J. R. Liu. In-car driver authentication using wireless sensing. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7595–7599, 2019.
- [35] F. Roider and K. Raab. Implementation and evaluation of peripheral light feedback for mid-air gesture interaction in the car. In *2018 14th International Conference on Intelligent Environments (IE)*, pages 87–90, 2018.
- [36] F. Roider, L. Reisig, and T. Gross. Just look : The benefits of gaze-activated voice input in the car. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI '18*, page 210–214, New York, NY, USA, 2018. Association for Computing Machinery.
- [37] E. Sabic, S. Mishler, J. Chen, and B. Hu. Recognition of car warnings : An analysis of various alert types. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '17*, page 2010–2016, New York, NY, USA, 2017. Association for Computing Machinery.
- [38] F. Sachara, T. Kopinski, A. Gepperth, and U. Handmann. Free-hand gesture recognition with 3d-cnns for in-car infotainment control in real-time. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 959–964, 2017.
- [39] H. Sandhaus and E. Hornecker. A woz study of feedforward information on an ambient display in autonomous cars. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings, UIST '18 Adjunct*, page 90–92, New York, NY, USA, 2018. Association for Computing Machinery.
- [40] G. M. Schmidbauer-Wolf and M. Guder. Usability and ux of a gaze interaction tool for front seat passengers : Evaluation of a gaze controlled optical feedback system in a car. In *Proceedings of Mensch Und Computer 2019, MuC'19*, page 677–681, New York, NY, USA, 2019. Association for Computing Machinery.
- [41] M. Schmidt, O. Bhandare, A. Prabhune, W. Minker, and S. Werner. Classifying cognitive load for a proactive in-car voice assistant. In *2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 9–16, 2020.

- [42] M. Schmidt, D. Helbig, O. Bhandare, D. Stier, W. Minker, and S. Werner. Assessing objective indicators of users' cognitive load during proactive in-car dialogs. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, UMAP'19 Adjunct, page 87–91, New York, NY, USA, 2019. Association for Computing Machinery.
- [43] W. F. Seng and L. Bin. Intelligent gestural control and ergonomic analysis in barrier-free car seat design. In *Proceedings of the 3rd International Conference on Industrial and Business Engineering*, ICIBE 2017, page 34–39, New York, NY, USA, 2017. Association for Computing Machinery.
- [44] G. Shakeri, J. H. Williamson, and S. Brewster. Novel multimodal feedback techniques for in-car mid-air gesture interaction. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '17, page 84–93, New York, NY, USA, 2017. Association for Computing Machinery.
- [45] G. Shakeri, J. H. Williamson, and S. Brewster. May the force be with you : Ultrasound haptic feedback for mid-air gesture interaction in cars. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '18, page 1–10, New York, NY, USA, 2018. Association for Computing Machinery.
- [46] G. Shakeri, J. H. Williamson, and S. A. Brewster. Bimodal feedback for in-car mid-air gesture interaction. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI '17, page 518–519, New York, NY, USA, 2017. Association for Computing Machinery.
- [47] K. A. Smith, C. Csech, D. Murdoch, and G. Shaker. Gesture recognition using mm-wave sensor for human-car interface. *IEEE Sensors Letters*, 2(2) :1–4, 2018.
- [48] G. Stevens, P. Bossauer, S. Vonholdt, and C. Pakusch. Using time and space efficiently in driverless cars : Findings of a co-design study. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, page 1–14, New York, NY, USA, 2019. Association for Computing Machinery.
- [49] A. Tamoto and K. Itou. Voice authentication by text dependent single utterance for in-car environment. In *Proceedings of the Tenth International Symposium on Information and Communication Technology*, SoICT 2019, page 336–341, New York, NY, USA, 2019. Association for Computing Machinery.
- [50] M. P. Tarvekar. Hand gesture recognition system for touch-less car interface using multiclass support vector machine. In *2018 Second In-*

*ternational Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 1929–1932, 2018.

- [51] S. Tatenno, Y. Zhu, and F. Meng. Hand gesture recognition system for in-car device control based on infrared array sensor. In *2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pages 701–706, 2019.
- [52] H. Tennent, D. Moore, and W. Ju. Character actor : Design and evaluation of expressive robot car seat motion. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(4), Jan. 2018.
- [53] A. Tewari, B. Taetz, F. Grandidier, and D. Stricker. [poster] a probabilistic combination of cnn and rnn estimates for hand gesture based interaction in car. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 1–6, 2017.
- [54] F. Weidner and W. Broll. Interact with your car : A user-elicited gesture set to inform future in-car user interfaces. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia, MUM '19*, New York, NY, USA, 2019. Association for Computing Machinery.
- [55] G. Young, H. Milne, D. Griffiths, E. Padfield, R. Blenkinsopp, and O. Georgiou. Designing mid-air haptic gesture controlled user interfaces for cars. *Proc. ACM Hum.-Comput. Interact.*, 4(EICS), June 2020.
- [56] A. Zhu and C. Zhang. The value and application of car head-up display in interactive design. In *Proceedings of the 2nd International Conference on Artificial Intelligence and Advanced Manufacture, AIAM2020*, page 478–483, New York, NY, USA, 2020. Association for Computing Machinery.